Time-Varying Frequency Analysis of Bat Echolocation Signals using Monte Carlo Methods

Sharad Nagappa

A thesis submitted for the degree of Doctor of Philosophy.
The University of Edinburgh.
2009

Abstract

Echolocation in bats is a subject that has received much attention over the last few decades. Bat echolocation calls have evolved over millions of years and can be regarded as well suited to the task of active target-detection. In analysing the time-frequency structure of bat calls, it is hoped that some insight can be gained into their capabilities and limitations.

Most analysis of calls is performed using non-parametric techniques such as the short time Fourier transform. The resulting time-frequency distributions are often ambiguous, leading to further uncertainty in any subsequent analysis which depends on the time-frequency distribution. There is thus a need to develop a method which allows improved time-frequency characterisation of bat echolocation calls.

The aim of this work is to develop a parametric approach for signal analysis, specifically taking into account the varied nature of bat echolocation calls in the signal model. A time-varying harmonic signal model with a polynomial chirp basis is used to track the instantaneous frequency components of the signal. The model is placed within a Bayesian context and a particle filter is used to implement the filter. Marginalisation of parameters is considered, leading to the development of a new marginalised particle filter (MPF) which is used to implement the algorithm. Efficient reversible jump moves are formulated for estimation of the unknown (and varying) number of frequency components and higher harmonics.

The algorithm is applied to the analysis of synthetic signals and the performance is compared with an existing algorithm in the literature which relies on the Rao-Blackwellised particle filter (RBPF) for online state estimation and a jump Markov system for estimation of the unknown number of harmonic components. A comparison of the relative complexity of the RBPF and the MPF is presented. Additionally, it is shown that the MPF-based algorithm performs no worse than the RBPF, and in some cases, better, for the test signals considered. Comparisons are also presented from various reversible jump sampling schemes for estimation of the time-varying number of tones and harmonics.

The algorithm is subsequently applied to the analysis of bat echolocation calls to establish the improvements obtained from the new algorithm. The calls considered are both amplitude and frequency modulated and are of varying durations. The calls are analysed using polynomial basis functions of different orders and the performance of these basis functions is compared. Inharmonicity, which is deviation of overtones away from integer multiples of the fundamental frequency, is examined in echolocation calls from several bat species. The results conclude with an application of the algorithm to the analysis of calls from the feeding buzz, a sequence of extremely short duration calls emitted at high pulse repetition frequency, where it is shown that reasonable time-frequency characterisation can be achieved for these calls.

Declaration of originality

I hereby declare that the research recorded in this thesis and the thesis itself was composed and originated entirely by myself in the Department of Electronics and Electrical Engineering at The University of Edinburgh.

All recordings of bat echolocation calls used herein were obtained from Dr Dean Waters, University of Leeds.

Sharad Nagappa

Acknowledgements

I would like to thank my supervisor Dr James Hopgood, whose guidance and support during the course of this work has proved invaluable. I would also like to thank Prof Steve McLaughlin for his motivation and encouragement throughout the last few years. Thanks are also due to Yannis Kopsinis who has always been helpful and encouraging, as well as Nicola Ferguson for her constant helpfulness. I would also like to acknowledge the support of those affiliated with the BIAS project for their contributions.

I would like to acknowledge the support of my colleagues, friends and fellow partners in crime, in particular, Nick Johnson, Mostafa Afgani, Gabriel Rilling, Sinan Sinanovic, Christine Evers, Mehrdad Yaghoobi-Vaighan and Donata Wasiuk.

Lastly, I would like to thank my parents and my brother, to whom I am indebted, for their patience, encouragement, love and support through the years.

Contents

List of figures

List of tables

Acronyms and abbreviations

AM	amplitude modulated
AR	autoregressive
CRLB	Cramer-Rao lower bound
CJM	combined jump move
CF	constant frequency
CDF	cumulative distribution function
CMF	cumulative mass function
CSW	cumulative sum of weights
EKF	extended Kalman filter
FM	frequency modulated
GMPHD	Gaussian mixture PHD
GHM	general harmonic model
GLM	general linear model
HMM	hidden Markov model
HAF	higher order ambiguity function
i.i.d.	independently and identically distributed
IF	instantaneous frequency
JMS	jump Markov system
LMS	least mean squares
LPFT	local polynomial Fourier transform

LPP	local polynomial periodogram
MPF	marginalised particle filter
MCMC	Markov chain Monte Carlo
MSE	mean squared error
MMSE	minimum mean squared error
MTJM	modified two-jump move
NLS	non-linear least squares
PDF	probability distribution function
PHD	probability hypothesis density
PRF	pulse repetition frequency
RBPF	Rao-Blackwellised particle filter
RJMCMC	reversible jump MCMC
SIR	sequential importance resampling
SMC	sequential Monte Carlo
STFT	short time Fourier transform
SNR	signal to noise ratio
SJM	single jump move
TVAR	time-varying autoregressive
TJM	two-jump move
UKF	unscented Kalman filter

Nomenclature

x	scalar
$\mathbf{x}$	vector
$\mathbf{x}_t$	vector at time t
$\mathbf{x}_{t_1:t_2}$	vectors $\mathbf{x}_t$ from time t_1 to t_2
$\mathbf{X}$	matrix
$\mathbf{X}^T$	matrix transpose
$\mathbf{X}^{-1}$	matrix inverse
$\mathcal{A}[s(t)]$	analytic signal of $s(t)$
$\mathcal{H}(s(t)]$	Hilbert transform of $s(t)$
$\mathcal{W}(t, \omega)$	Wigner distribution at time t and frequency ω
$s(t) \rightleftharpoons S(\omega)$	Fourier transform pairs
$<\mathbf{x}>$	mean of $\mathbf{x}$
$\delta(t)$	Dirac delta function
$\mathfrak{R}(c)$	real part of complex term c
$\mathfrak{I}(c)$	imaginary part of complex term c
$\mathbb{I}_Q$	$Q \times Q$ identity matrix
$p(\cdot)$	probability density function
$P(\cdot\|\cdot)$	conditional probability density function
$\mathcal{U}[a, b]$	uniform distribution on the interval $[a, b]$
$\mathcal{U}_d[a, b]$	discrete uniform distribution on the interval $[a, b]$
$\mathcal{N}(\mu, \Sigma)$	Gaussian distribution with mean μ and covariance matrix Σ
$\mathcal{IG}(\alpha, \beta)$	inverse gamma distribution with scale parameter α and shape parameter β
$\sigma^2_{(x),t}$	variance of the parameter x at time t
$\min\{\}$	minimum value
$\mathcal{O}(\cdot)$	computational complexity using big O notation
$\|\cdot\|$	Frobenius norm
$x \sim p(\cdot)$	x is drawn from the distribution $p(\cdot)$

Chapter 1
Introduction

Bats are of interest due to their almost exclusive use of sonar for navigation and hunting. They are able to detect and capture prey, consisting mostly of insects, even while hunting in the close vicinity of dense foliage. The echoes from insects are often buried under the echoes from vegetation, which is referred to as *clutter*. Research on bats has covered such aspects as echolocation call structure [1–5], ear construction and neural processing mechanisms (see Part Two of [1]). It is analysis of call structure, however, that is of particular interest in this work.

Most analysis of bat echolocation calls relies on spectrogram-based analysis. The use of alternative techniques for call analysis, for example using the Wigner Ville distribution or other high resolution spectrum analysis methods, is uncommon [6]. Kopsinis *et al*. [6] provide a comparison of different time-varying frequency estimation techniques and illustrate their application to bat echolocation calls.

Previous studies have predominantly relied on qualitative comparisons between calls, in part due to the difficulty in arriving at a quantitative description of the calls. The aim of the work undertaken here is to provide a means to facilitate quantitative descriptions of echolocation calls. This is achieved through estimation of the time-varying frequencies which constitute the call. Availability of the time-frequency structure of the calls can provide a stepping stone to further analysis.

The following section summarises the motivation for the approach adopted here. This is followed by an overview of the layout of the rest of this thesis.

1.1 Motivation for the Approach

Analysis of bat echolocation calls can be performed using a variety of non-parametric time-frequency estimation methods. The resulting time-frequency distributions can sometimes be ambiguous, and furthermore, it is necessary to further process these distributions to extract the time-varying frequencies associated with the call, making such an approach less than ideal.

Parametric approaches to the time-frequency estimation problem can incorporate prior knowledge of the signal structure within a model. When dealing with bat echolocation calls, it can be difficult to fit a model to the entire call and subsequently estimate the parameters. An alternative approach is to sequentially estimate the call frequencies using short segments of the call. A time-varying sinusoidal model provides a means of modelling signal structure as it evolves over time. However, the sinusoidal model is unsuitable for the analysis of frequency modulated signals and bat echolocation calls can possess high degrees of frequency modulation.

The approach adopted here is to use a harmonic model with a polynomial chirp basis, thus allowing characterisation of frequency modulation within the signal. The harmonic observation model is highly non-linear which presents a difficult parameter estimation exercise. When considered within a sequential context, however, this problem can be solved using particle filtering methods.

When the harmonic model is used in a batch offline scenario, it can be shown that it is possible to marginalise parameters from the posterior distribution which simplifies the estimation problem. In this thesis it will be shown that this marginalisation technique can be extended to the particle filter framework and provides an alternative to the well established Rao-Blackwellised particle filter (RBPF).

In order to deal with the time-varying number of harmonics present in echolocation calls, a reversible jump sampler can be used. Reversible jump samplers have been developed elsewhere for determining an unknown number of tones and/or harmonics in a signal. When used in a sequential framework, the reversible jump sampler results in a large computational burden. Consequently, efficient reversible jump schemes are developed in this work which do not reduce the quality of the resulting frequency estimates.

Using the algorithms developed herein, analysis of bat echolocation calls is considered and improvements from the methods developed here are illustrated.

1.2 Thesis Layout

Chapter 2 serves to introduce the topic of time-frequency analysis. Beginning with a discussion on the concept of instantaneous frequency, the chapter goes on to discuss various parametric and non-parametric methods which have been developed for time-frequency analysis. In particular,

significant attention is devoted to a discussion of the short time Fourier transform (STFT), the Wigner distribution, as well as the implication of limited time-frequency resolution as governed by the Heisenberg-Gabor uncertainty principle.

The second chapter concludes with an overview of bat echolocation calls. The aim is to stress the variability and ambiguity associated with these calls, which in turn provides the motivation and justification for the assumptions and methods developed for call analysis in this thesis.

Chapter 3 continues with the discussion of time-frequency analysis, but is restricted to those methods which are based on sequential Monte Carlo (SMC) methods. To this end, the chapter begins with a discussion of particle filtering methods. The chapter also introduces parameter marginalisation as achieved using the RBPF, as well as describing the jump Markov system (JMS) and reversible jump MCMC (RJMCMC) methods for estimating an unknown model order. The chapter also considers analysis of bat echolocation calls to illustrate the shortcomings of some of the approaches listed.

Chapter 4 introduces the signal model which will be used in the subsequent analysis. The signal model draws on harmonic models described in various existing literature. These models are usually adapted to various specific situations, and the model described in the chapter serves to bring these variations together through a general model. Also considered here is marginalisation of model parameters which finds application in the subsquently derived marginalised particle filter (MPF).

Chapter 5 develops a new algorithm for sequential state estimation called the MPF which is subsequently applied to the problem of sequential frequency estimation. The chapter compares the MPF, as well as the sequential frequency estimation algorithm, with existing methods to illustrate its benefits.

The estimation algorithm developed in Chapter 5 relies on the use of reversible jump moves to detect the number of frequency components, resulting in an algorithm with high computational complexity. Chapter 6 considers alternative reversible jump schemes to reduce the complexity of the algorithm without sacrificing the efficacy of the moves. A simplifying assumption is made in the model to develop new computationally less demanding schemes.

Analysis of bat echolocation calls is considered in Chapter 7. The chapter examines the impact of using different orders of the polynomial basis during analysis of calls. Additionally, the

presence of inharmonicity in bat calls from several different species is examined. The chapter concludes with an analysis of calls from the *feeding buzz* of an echolocation call sequence. Hyper-parameter tuning is discussed to improve the quality of the frequency estimates obtained from the algorithm.

The concluding chapter of this work examines the achievements of the work as well as its limitations. Further research which builds on the work carried out here is considered to improve both the method as well as the analysis of echolocation calls.

Chapter 2
Time-Frequency Analysis

Introduction

As described in the previous chapter, the aim of this work is to analyse the time-frequency structure of bat echolocation calls. Before looking at bat echolocation calls themselves, this chapter summarises various parametric and non-parametric methods for time-frequency analysis. The advantage of using non-parametric methods is that very few assumptions are made about the signal of interest. Application of any method will always provide a result, however, it is up to the user to apply the method in a consistent and sensible manner so that the result is meaningful. As such, non-parametric methods can often be used to gain insight into the structure of the signal, subsequently leading to the design and refinement of a parametric signal model [7, 8].

The subject of interest being the time-varying frequency content of a signal, the following section introduces the subject of the instantaneous frequency (IF) and its implications for mono- and multi-component signals. Following this discussion, non-parametric linear and bilinear time-frequency distributions are examined. While these various time-frequency distributions provide a visual representation of the time-varying spectrum, determination of precise frequency components in the signal requires post-processing of these representations.

As an alternative, parametric approaches to time-frequency analysis can be adopted. Using these methods, a direct estimate of the (time-varying) frequency content can be obtained, thus removing the need for any form of post-processing. Some parametric methods are covered in the concluding section of this chapter, while the following chapter deals with sequential Monte Carlo (SMC) based methods for time-frequency analysis. As a specific case of parametric time-frequency analysis, the section on parametric approaches concludes with an examination of chirp parameter estimation. In some cases, the structure of the chirp is known beforehand, and the parameters of the signal can be estimated via the chirp parameters.

The chapter concludes with an introduction to bat echolocation calls. The section describes trends which are seen in call variation as well as describes various chirp models which have

been used for fitting the calls. The methods described in this chapter can be used to learn about the structure of echolocation calls. However, as discussed later in this chapter, there are certain limitations associated with these methods in their analysis of echolocation calls. Consequently, a new SMC approach is developed in Chapter 5 for this purpose.

2.1 The Concept of Instantaneous Frequency

Before discussing methods for time-frequency analysis in the following sections, the topic of instantaneous frequency (IF) is addressed. The term IF is often used ambiguously or incorrectly and the aim of this section is to clarify what exactly is meant by *instantaneous frequency*.

The IF [7, 9–11] of a signal is of interest in many signal analysis applications, for example, in seismic, radar and biomedical processing. The IF of signals in these cases may provide information regarding the target of interest, for example, in terms of structure of rocks or target velocity (as a Doppler shift in the radar return). In the case of echolocation calls, the IF relates to the signal structure and can be used to analyse the time-frequency structure of calls.

Consider a monocomponent signal, *i.e.*, a signal with only a single frequency component. For such a monocomponent signal, the IF is a physically meaningful quantity, corresponding to the time-varying location of the signal's spectral peak. Under such circumstances, the IF may be understood as the "frequency of a sine wave which locally fits the signal under analysis" [9]. When dealing with multicomponent signals, however, the IF does not reflect the spectral content of the signal and can, in certain cases, be negative or consist of frequencies which do not occur in the spectrum [7, 9, 12]. In such cases, the IF of each *component* may be of interest, rather than the IF of the signal as a whole.

The IF at any point in time may be defined as the derivative of the phase of the signal at that time, such that its average over time yields the average frequency of the signal [7]. Consider a real signal of the form:

$$s(t) = a(t)\cos\left[\phi(t)\right] \tag{2.1}$$

where $a(t)$ and $\phi(t)$ represent the time-varying amplitude and phase of the signal $s(t)$. For such a signal, the instantaneous frequency $f_i(t)$ can be written as [9]:

$$f_i(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt} \tag{2.2}$$

In the case of a complex signal of the form:

$$z(t) = s_r(t) + js_i(t) = a(t)e^{j\phi(t)} \tag{2.3}$$

where $s_r(t)$ and $s_i(t)$ represent the real and imaginary parts of the signal. The magnitude, phase, and IF of the complex signal, $a(t)$, $\phi(t)$, and $f_i(t)$ respectively are given by [7][1]:

$$a(t) = \sqrt{s_r(t)^2 + s_i(t)^2} \tag{2.4}$$

$$\phi(t) = \arctan\left(\frac{s_i(t)}{s_r(t)}\right) \tag{2.5}$$

$$f_i(t) = \frac{d\phi(t)}{dt} = \frac{s_i'(t)s_r(t) - s_r'(t)s_i(t)}{a(t)^2}. \tag{2.6}$$

If the IF is defined as the derivative of the phase, the above definition implies that the IF of a real signal is zero, in apparent contrast to the definition of equation (2.2), a notion pointed out as absurd by Cohen [7]. In order to reconcile these two definitions, the real signal is rewritten in complex form such that the amplitude and phase are preserved. The analytic signal offers one means of representing the real signal as its complex counterpart (another method being the quadrature method; see [7] for a discussion). Apart from resolving the apparent differences between the two definitions for the IF, there are other advantages to converting a real signal to its analytic signal counterpart [7].

Consider the Fourier transform pairs $s(t) \rightleftharpoons S(\omega)$. For a real signal, the spectrum is symmetric about the origin and satisfies the condition $S(-\omega) = S^\star(\omega)$. Consequently, an estimate of the average frequency of such a signal, computed as the weighted average of the spectrum, is always zero. Further, the standard deviation of the frequency will be approximately the spread between the spectral components on either side of the origin. These misleading estimates may be circumvented by only considering the spectrum in the range $[0, \infty]$. Alternatively, it is possible to redefine the signal such that the spectrum remains unchanged for the positive frequencies, but it has no negative frequencies and the frequency averages can be computed directly from the signal.

[1] $\frac{d}{dx}\arctan(x) = \frac{1}{1+x^2}$; $\frac{d}{dt}\arctan\left(\frac{y}{x}\right) = \frac{y'x - x'y}{x^2} \times \frac{1}{1+\frac{y^2}{x^2}}$

2.1.1 The Analytic Signal

The analytic signal, as mentioned in the previous section, provides an equivalent representation of a real signal in complex form while preserving the amplitude and phase of the real signal. The analytic signal is written as [7, 9]:

$$\mathcal{A}[s(t)] = z(t) = s(t) + \mathcal{H}[s(t)] = a(t)e^{j\phi(t)} \tag{2.7}$$

where $s(t)$ denotes the real signal and $\mathcal{H}[s(t)]$ denotes the Hilbert transform operator on the signal $s(t)$. The Hilbert transform is defined as [9]:

$$\mathcal{H}[s(t)] = \text{p.v.} \int_{-infty}^{+\infty} \frac{s(t-\tau)}{\pi\tau} d\tau \tag{2.8}$$

where p.v. denotes the Cauchy principal value of the integral.

The result of the analytic signal procedure can be interpreted as placing the low frequency content of the signal in the amplitude modulated term $a(t)$, while placing the high frequency content in the phase modulated term $e^{j\phi(t)}$ [7].

Cohen [7] notes that there are certain paradoxes in defining the IF as the derivative of the phase of the analytic signal. There is also the additional issue of whether this definition of the IF represents anything meaningful and Cohen lists the following issues as being counter-intuitive to the above definition of the IF [7]:

1. The IF may not be one of the frequencies present in the spectrum of the signal. Consider a signal consisting of a line spectrum with a few well defined frequencies. The IF of such a signal may be both continuous, as well as range over an infinite number of values.

2. The spectrum of the analytic signal is defined as non-negative, however, the IF of the analytic signal may go negative.

3. The IF may go outside of the bandwidth of a bandlimited signal.

4. The IF indicates the frequency of a signal at a particular instant of time, and, as such, should not depend on past or future behaviour of the signal. However, construction of the analytic signal requires knowledge of the signal for all time since the integral in the Hilbert transform is carried out over $[-\infty, +\infty]$.

Cohen also notes that some of these issues are resolved by considering the IF as a conditional mean frequency, which is the average frequency of the signal at each time instant [7, 11, 13, 14].

2.1.2 The Uncertainty Principle

The Heisenberg-Gabor uncertainty principle (or simply, the uncertainty principle) [7, 8] places a fundamental limitation on the estimation of time-frequency distributions. It states that it is impossible to obtain arbitrarily fine resolution simultaneously in terms of both time and frequency. In the words of Skolnik [15]:

> "It states the well-known mathematical fact that a narrow waveform yields a wide spectrum and a wide waveform yields a narrow spectrum and both the time waveform and frequency spectrum cannot be made arbitrarily small simultaneously."

Mathematically, the uncertainty principle can be expressed as follows. Consider the Fourier transform pairs $s(t) \rightleftharpoons S(\omega)$. If the energy density of the signal is defined as $|s(t)|^2$, where $|\cdot|$ denotes the magnitude, the *duration* of the signal is defined as [7]:

$$T^2 = \sigma_t^2 = \int (t - \langle t \rangle)^2 |s(t)|^2 dt \tag{2.9}$$

$$= \langle t^2 \rangle - \langle t \rangle^2 \tag{2.10}$$

where:

$$\langle t \rangle = \int t |s(t)|^2 dt \tag{2.11}$$

$$\langle t^2 \rangle = \int t^2 |s(t)|^2 dt \tag{2.12}$$

and $\langle t \rangle$ denotes the mean time of the signal. Similarly, the energy density spectrum is given by $|S(\omega)|^2$ and the (root mean square) bandwidth is denoted by:

$$B^2 = \sigma_\omega^2 = \int (\omega - \langle \omega \rangle)^2 |S(\omega)|^2 d\omega \tag{2.13}$$

$$= \langle \omega^2 \rangle - \langle \omega \rangle^2 \tag{2.14}$$

where:

$$\langle \omega \rangle = \int \omega |S(\omega)|^2 d\omega \tag{2.15}$$

$$\langle \omega^2 \rangle = \int \omega^2 |S(\omega)|^2 d\omega \tag{2.16}$$

and $\langle \omega \rangle$ denotes the average frequency.

The average time and frequency denote the concentration of the signal in the time and frequency distributions respectively while the variance on these values, denoted by T^2 and B^2, indicates the spread around the mean. Consequently, T^2 and B^2 are measures of the signal's duration and bandwidth respectively [7].

Using the above definitions, the uncertainty principle can be written as [7]:

$$TB = \Delta t \Delta \omega \geq \frac{1}{2} \tag{2.17}$$

where $\Delta t = T$ and $\Delta \omega = B$.

Thus, as described by Cohen [7], if the density of the signal in time is given by $|s(t)|^2$ and the frequency density is $|S(\omega)|^2$, then both distributions cannot be made narrow simultaneously. If the time density is made narrow, this results in a broad frequency density (and vice versa) which is reflected by the relation between the terms T and B in equation (2.17).

2.1.3 Time-Frequency Distributions

Before addressing the various means of calculating joint time-frequency distributions, some properties of time-frequency distributions are discussed here.

Let a joint time-frequency density be denoted by $P(t, \omega)$ at time t and frequency ω. The energy density and the energy density spectrum are obtained from the joint density as marginals [7, 8]:

$$P(t) = \int P(t, \omega)\, d\omega = |s(t)|^2 \tag{2.18}$$

$$P(\omega) = \int P(t, \omega)\, dt = |S(\omega)|^2 \tag{2.19}$$

and the total energy is given by:

$$E = \iint P(t,\omega)\, dt\, d\omega = \int |s(t)|^2\, dt = \int |S(\omega)|^2\, d\omega \tag{2.20}$$

It is possible for a joint time-frequency distribution to satisfy the total energy requirement of equation (2.19) without satisfying the marginals in equations (2.18) and (2.19), the total energy requirement being a weak one [7]. The short time Fourier transform (STFT) is one time-frequency distribution that does not satisfy the marginals, however, it is able to produce a reasonably good time-frequency representation.

2.2 Non-Parametric Time-Frequency Analysis

The previous section discussed the concept of instantaneous frequency as well as introduced the analytic signal which ensures that the spectrum contains no negative frequencies. This section introduces non-parametric time-frequency analysis methods and examines their properties. These methods can be used to generate time-frequency distributions which offer a means of estimating the time-varying spectrum and IF of a signal.

Non-parametric methods seek to decompose the signal onto a *maximal* set of elementary basis functions. The projection of the signal on to each function is a measure of the contribution of that basis in the decomposition. A decomposition of such a form can be used to generate a distribution indicating the spectrum of the signal. The Fourier transform, for example, uses a set of sinusoids as the basis and the projection of the signal on to each sinusoid is a measure of that frequency in the signal as obtained from the squared magnitude of the Fourier coefficients.

This section provides a summary of time-frequency analysis using the STFT, wavelets, and concludes with a discussion of the Wigner distributions and other bilinear time-frequency distributions. Following this, Section 2.3 examines parametric approaches to time-frequency estimation using online state estimation, time-varying autoregressive (TVAR) methods and chirp estimation.

2.2.1 Short Time Fourier Transform

The STFT [7, 8, 16] offers a means to extend the Fourier transform to the study of time-varying signals. This is performed by breaking the signal into short segments and applying the Fourier

transform to each segment to determine its frequency spectrum.

The STFT uses a window to localise the signal in time. Mathematically, for a signal $s(t)$ and a window $h(t)$ the localised signal is obtained as [7]:

$$s_t(\tau) = s(\tau)h(\tau - t). \tag{2.21}$$

The corresponding spectrum is given by:

$$S_t(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-j\omega\tau} s_t(\tau) d\tau \tag{2.22}$$

and the STFT spectrogram of the signal is then given by $|S_t(\omega)|^2$.

The window $h(t)$ must be chosen such that it possesses finite support in the time domain. Additionally, the above formulation does not discriminate between the window and the signal which can be interchanged without any consequence. Thus, as noted by Cohen [7], one has to be "careful that we are not using the signal to study the window". The window may be chosen to highlight certain properties of the signal. For example, a rectangular window has poor side-lobe suppression, so that it will not be possible to distinguish multiple closely spaced frequency components which appear within the side-lobes of a stronger frequency component.

Changing the width of the window can be used to obtain better localisation of signal characteristics in either the time or the frequency domain. A short window is able to better localise characteristics in the time domain. In particular, as the window is narrowed such that it becomes the delta function, $h(t) \to \delta(t)$, the estimated average frequency approaches the derivative of the phase of the signal, or the IF, $\langle\omega\rangle \to \frac{d\phi(t)}{dt}$ [7]. However, as the window narrows, the standard deviation, $\sigma_{\omega|t}$, of the instantaneous frequency estimate at time t broadens, and as the estimate approaches the true IF, it becomes impossible to accurately measure it [7]:

$$\sigma_{\omega|t} \to \infty \quad \text{as } h(t) \to \delta(t). \tag{2.23}$$

The use of shorter segments leads to better localisation of the signal characteristics in time, however, as the segments get shorter, the STFT spectrum loses meaning at some point, bearing no relation to the spectrum of the original signal [7] due to the increase in variance. This limitation is frequently referred to as the time-frequency resolution trade-off. Despite this limitation, the STFT provides good time-frequency representations for many signals, although, it may not

always provide the clearest representation [7].

As mentioned earlier, the STFT does not satisfy the marginals of equations (2.18) and (2.19). Consider the Fourier transform pairs of the signal, $s(t) \rightleftharpoons S(\omega)$, and the window $h(t) \rightleftharpoons H(\omega)$:

$$s(t) = A(t)e^{j\phi(t)} \rightleftharpoons B(\omega)e^{j\psi(\omega)} = S(\omega) \tag{2.24}$$

$$h(t) = A_h(t)e^{j\phi_h(t)} \rightleftharpoons B_h(\omega)e^{j\psi_H(\omega)} = H(\omega). \tag{2.25}$$

Integrating the joint distribution $S_t(\omega)$ in equation (2.22) over the frequency term yields the time marginal [7]:

$$\begin{aligned} P(t) &= \int |S_t(\omega)|^2 \, d\omega \\ &= \int A^2(\tau)A_h^2(\tau - t)d\tau. \end{aligned} \tag{2.26}$$

Similarly, the frequency marginal is obtained as an integral over time:

$$P(\omega) = \int B^2(\omega')B_h^2(\omega - \omega')d\omega'. \tag{2.27}$$

The marginals evaluated above do not satisfy the correct marginals given in equations (2.18) and (2.19) [7]:

$$P(t) \neq A^2(t) = |s(t)|^2 \tag{2.28}$$

$$P(\omega) \neq B^2(\omega) = |S(\omega)|^2 . \tag{2.29}$$

The average frequency $\langle\omega\rangle$ given by equation (2.15) cannot be calculated from the STFT since $|S(\omega)|^2$ cannot be evaluated as a marginal of the joint time-frequency distribution. Cohen notes that in calculating the STFT, the energy distributions of the signal and the window get scrambled together, such that the average frequency cannot be measured accurately, although it may be possible to make a good approximation in some cases [7]. Section 2.2.3 will discuss time-frequency distributions that satisfy the marginals, thus allowing the average time and frequency to be measured.

2.2.1.1 Local Polynomial Approximations

The underlying assumption in using the STFT is that a sinusoidal basis provides an accurate decomposition of the signal. If the signals being analysed are non-stationary within the window, the resulting spectrogram may be misleading. Non-stationary, for example, may denote the presence of frequency or amplitude modulation of the signal which is not taken into account in the model assumptions.

Modifications to the STFT have been suggested which rely on characterising the local frequency content using a polynomial basis rather than a constant frequency basis [17–19]. Such an approach can be seen as modelling the non-stationarity present within the window to achieve a better time-frequency representation. Using a local polynomial frequency approximation however, can be computationally expensive.

Flandrin [16] provides an overview of chirp analysis methods using modifications to the Fourier transform as well as wavelet decompositions. Bretthorst [17] and Jaynes [18] define a *chirpogram* which is the Fourier transform calculated over the frequency and chirp rate parameter pair. This is extended further by Katkovnik [19] to use an arbitrary m^{th} order polynomial for the basis giving rise to the local polynomial Fourier transform (LPFT) and the local polynomial periodogram (LPP). The expression for the time-varying spectrum of the STFT is then adapted as:

$$S_t(\tilde{\omega}) = \int e^{-j\theta(\tau,\tilde{\omega})} s_t(\tau) d\tau, \quad \tilde{\omega} = [\omega_1, \ldots, \omega_m]^T \in \mathcal{R}^m \tag{2.30a}$$

$$\theta(\tau,\tilde{\omega}) = \tau\omega_1 + \frac{\tau^2}{2}\omega_2 + \cdots + \frac{\tau^m}{m!}\omega_m \tag{2.30b}$$

where $\tilde{\omega}$ represents the coefficients of the m^{th} order polynomial, and $\theta(\tau,\tilde{\omega})$ is a function of the time position in the window τ and $\tilde{\omega}$ such that $e^{-j\theta(\tau,\tilde{\omega})}$ forms a complex polynomial chirp. Using the LPFT, the LPP is defined as $|S_t(\tilde{\omega})|^2$.

2.2.1.2 Reassigned Spectrogram

The reassignment method [8, 20] is a means of improving the quality of the STFT. The reassignment method has been generalised to any bilinear time-frequency representation in [20] and applied to the LPFT in [21] (bilinear time-frequency distributions are discussed in Section 2.2.3).

The reassignment method as applied to the spectrogram works by analysing the phase term in the spectrogram and "reassigns energy away from [the] sampling lattice point (t, ω) to the centre of gravity of the windowed energy $(\hat{t}, \hat{\omega})$" [22]:

$$\hat{t} = t - \Re\left(\frac{STFT_{th}}{STFT_h}\right) \tag{2.31a}$$

$$\hat{\omega} = \omega + \Im\left(\frac{STFT_{dh}}{STFT_h}\right) \tag{2.31b}$$

where $\Re(\cdot)$ and $\Im(\cdot)$ respectively denote the real and imaginary part of their arguments. $STFT_h$ indicates evaluation of the STFT using the window h as described by equation (2.22). $STFT_{th}$ denotes evaluation of the STFT using a window weighted by a time-ramp [23]:

$$STFT_{th} = \frac{1}{\sqrt{2\pi}} \int e^{-j\omega\tau} \tau s_t(\tau) d\tau \tag{2.32}$$

and $STFT_{dh}$ denotes the use of a window weighted by the time-derivative of the window. Weighting by the time-derivative of the window is equivalent to weighting the Fourier transform with a frequency ramp [22], so that $STFT_{dh}$ can be evaluated as [23]:

$$STFT_{dh} = \frac{1}{\sqrt{2\pi}} \int e^{-j\omega\tau} \frac{dh(t)}{dt} s_t(\tau) d\tau \tag{2.33}$$

$$= \frac{1}{\sqrt{2\pi}} \int e^{-j\nu t} S(\nu) H(\nu - \omega) d\nu \tag{2.34}$$

$$\tag{2.35}$$

Thus, the contribution from the term $S_t(\omega)$ is reassigned to the location $(\hat{t}, \hat{\omega})$, and the resulting energy at the point $(\hat{t}, \hat{\omega})$ is obtained as the sum of all contributions which are reassigned to that point.

Through this process, the spread of the components in the STFT is reduced, and the components can be highly localised in the time-frequency plane. Reassigned spectrograms have been used, for example, for musical transcription in [22] and to analyse gravitational waves at low signal to noise ratio (SNR) in [24] where Duvaut *et al.* showed that the reassignment method works as well as a Bayesian solution.

2.2.2 Wavelets

Wavelets [8,25–27] offer an alternative means for studying non-stationary signals through time-scale analysis. Wavelet analysis relies on a decomposition of the signal on to a set of zero-mean basis functions which are localised in both time and frequency. The wavelets are derived from translations and dilations of the mother wavelet, allowing analysis of the signal at different instants and on different scales. The dilations result in altering the frequency spread of the wavelet, while the zero-mean property ensures that the frequency position is not affected [26].

Wavelets thus examine large scale features over longer durations of time and small scale features over short durations of time. Since large scale features are estimated over longer time periods, the resulting scalogram is not evaluated over a rectangular grid structure as in the case of the STFT, but over a dyadic mesh [8].

Wavelets are mentioned here as an alternative to the STFT, however, their properties and application are not considered. The following section expands on time-frequency distributions by introducing bilinear time-frequency distributions, and, in particular, examining the Wigner distribution and its properties.

2.2.3 Bilinear Time-Frequency Distributions

A general class of time-frequency distributions can be written in the form [7, 8]:

$$C(t,\omega) = \frac{1}{4\pi^2} \iiint s^\star \left(u - \frac{1}{2}\tau\right) s\left(u + \frac{1}{2}\tau\right) \phi(\theta,\tau) e^{-j\theta t - j\tau\omega + j\theta u} du d\tau d\theta \qquad (2.36)$$

where $s^\star(t)$ denotes the complex conjugate and $\phi(\theta,\tau)$ is a two dimensional function called the kernel. These *bilinear* distributions, are called so due to the presence of the signal twice in the calculation as $s^\star(u - \frac{1}{2}\tau)s(u + \frac{1}{2}\tau)$.

In this section, the Wigner distribution is discussed in detail. The limitations and drawbacks of the Wigner distribution have lead to the development of other distributions which are summarised in the conclusion of this section.

Setting the kernel, $\phi(\theta, \tau)$ in equation (2.36) to 1 results in the Wigner distribution:

$$\mathcal{W}(t,\omega) = \frac{1}{2\pi}\int s^{\star}\left(t-\frac{1}{2}\tau\right)s\left(t+\frac{1}{2}\tau\right)e^{-j\tau\omega}d\tau \tag{2.37}$$

$$= \frac{1}{2\pi}\int S^{\star}\left(\omega+\frac{1}{2}\theta\right)S\left(\omega-\frac{1}{2}\theta\right)e^{-jt\theta}d\theta \tag{2.38}$$

Although the Wigner distribution is shown here to belong to the class of time-frequency distributions given by equation (2.36), it was the Wigner distribution that was derived first, with the general class of time-frequency distributions derived later. The Wigner distribution is particularly important in that while improving upon certain aspects of the STFT, it is also saddled with its own limitations. Other time-frequency distributions were developed with the intent of improving upon the Wigner distribution.

Two important drawbacks are associated with the Wigner distributions. The first is its non-positivity. A bilinear distribution that satisfies the marginals cannot be positive throughout the time-frequency plane, and thus, the Wigner distribution always goes negative in some region [7]. While one exception to this occurs when dealing a zero signal $s(t) = 0$, for most signals of interest, the Wigner distribution may be expected to go negative.

The other drawback of the Wigner distribution is its inability to deal with multicomponent signals. For a multicomponent signal, the Wigner distribution places values in between the components in the time-frequency plane, giving rise to what are referred to as *cross terms* [7,8].

Consider a signal of the form:

$$s(t) = s_1(t) + s_2(t) \tag{2.39}$$

The Wigner distribution of this multicomponent signal is obtained as [7]:

$$\mathcal{W}(t,\omega) = \mathcal{W}_{11}(t,\omega) + \mathcal{W}_{22}(t,\omega) + 2\Re\{\mathcal{W}_{12}(t,\omega)\} \tag{2.40}$$

where

$$\mathcal{W}_{ij}(t,\omega) \triangleq \frac{1}{2\pi}\int s_i^{\star}\left(t-\frac{1}{2}\tau\right)s_j\left(t+\frac{1}{2}\tau\right)e^{-j\tau\omega}d\tau \tag{2.41}$$

and $\Re\{\mathcal{W}_{12}(t,\omega)\}$ represents the real part of the complex term.

Thus, the Wigner distribution of the sum of two signals is not merely the sum of their respective Wigner distributions, but also includes a cross term, or interference term [7], which appears

between the distributions described by $\mathcal{W}_{11}(t,\omega)$ and $\mathcal{W}_{22}(t,\omega)$ in the time-frequency plane. The presence of interference terms in the time-frequency distribution limits the usability of the Wigner distribution for analysis of multicomponent signals.

In addition to the above issues, the Wigner distribution is also highly non-local [7]. For example, noise which may be present in only one part of the signal affects the entire time-frequency distribution since all values from the past and future of the signal contribute to the distribution at time t. Thus, noise which appears in only one part of the signal will contribute even at other parts where there is no noise, resulting in a noisy time-frequency distribution.

The expression for the Wigner distribution in equation (2.37) requires an integration over all time, $[-\infty, +\infty]$. In practice, performing the integration over these limits may not be possible, nor is this desirable, as it can lead to a noisy representation as discussed previously. These limitations lead to the pseudo Wigner distribution, which seeks to make the distribution more local, while also mitigating the effect of cross terms in the distribution.

The pseudo Wigner distribution achieves this by introducing a windowing operation into the representation which emphasises the signal around the time of interest, while suppressing it elsewhere. The pseudo Wigner can then be written as [7]:

$$\mathcal{W}_{PS}(t,\omega) = \int h(\tau) s^{\star}\left(t - \frac{1}{2}\tau\right) s\left(t + \frac{1}{2}\tau\right) e^{-j\omega\tau} d\tau. \tag{2.42}$$

While the windowing operation serves to make the distribution more local, as well as suppresses the cross terms, it also destroys certain properties of the Wigner distributions. For example, the time and frequency marginals in equations (2.18) and (2.19) do not hold for the pseudo Wigner distribution [7].

One of the main reasons for modifying the Wigner distribution has been to achieve a positive distribution. It is possible to smooth the Wigner distribution by convolving it with a suitable smoothing function. As such, it is possible to convert the Wigner distribution into the STFT by convolving, in both time and frequency, the Wigner distribution of the signal with the Wigner distribution of a window. In fact, for all time-frequency distributions which can be written in the form of equation (2.36), each distribution can be viewed as smoothed versions of any other distribution. Wigner showed that it is not possible for a bilinear distribution to be positive and still satisfy the time and frequency marginals. It is possible to generate positive distributions

which satisfy the marginals, however, these distributions are not bilinear.

Another reason for modifying the Wigner distribution is to reduce the presence of cross terms in the distribution. Time-frequency distributions with reduced interference or cross terms can be created by careful selection of the kernel. The condition on the kernel to minimise the cross terms is [7]:

$$\phi(\theta, \tau) \ll 1, \quad \text{for } \theta \cdot \tau \gg 0 \tag{2.43}$$

which implies that the kernel is localised about the axes of θ and τ. For the Wigner distribution, the kernel is uniform in the time-frequency plane, $\phi(\theta, \tau) = 1, \forall \theta, \tau$ and thus suffers from cross terms. In contrast, the STFT kernel depends on the choice of window and thus sometimes exhibits good behaviour with regards to the cross terms.

Time-frequency distributions are a tool to examine the time-varying spectrum of a signal. They do not, however, make assumptions about the signal structure. In applying a particular decomposition, it is assumed that the decomposition is suitable and meaningful. The following section examines various parametric time-frequency analysis methods. These methods rely on the use of a suitable model and are likely to yield misleading results for cases when the signal does not reasonably match the model used.

2.3 Parametric Time-Frequency Analysis

Non-parametric time-frequency analysis relies on a decomposition of the signal onto a maximal set of basis functions. In contrast, parametric methods attempt to account for the manner in which the signal is constructed by using a signal model. Rather than providing a time-frequency distribution over all possible frequencies, they attempt to identify only those locations in the time-frequency plane which correspond to frequencies in the signal.

Parametric methods are useful when it is possible to provide a suitable model for the signal. Using such a model, it is possible to characterise a signal using a small number of parameters. In the context of time-frequency analysis, the parameter of interest could be the instantaneous frequency of a signal.

This section examines several methods for parametric estimation of the instantaneous frequency and/or spectrum of a signal. The methods may be online, using the Kalman filter or SMC methods for example, or batch methods. In the case of batch methods, the model needs to

account for the long-term evolution of the signal, and knowledge of the signal's modulation pattern is necessary.

Since these methods rely on the use of a signal model, it is essential that the model and the observation should be in agreement to some extent. For example, a harmonic model says nothing about an impulse type signal. The applicability of parametric methods to a particular signal thus depends on both the signal as well as the signal model.

The following discussion first examines batch approaches to time-frequency analysis before considering the online case.

2.3.1 Batch Offline Methods for Frequency and Time-Frequency Estimation

A parametric approach to time-frequency estimation can be adopted in cases where the underlying frequency and amplitude modulation can be modelled. Batch offline methods, in particular, find widespread application to frequency analysis of stationary signals. Their extension to non-stationary signals requires knowledge of the signal structure. As a result, while this makes them powerful tools for signal analysis, they may not be widely applicable. This section focuses on a few Bayesian solutions to the problem of spectrum analysis. Chapters 4 and 5 build on these methods and extend their application to non-stationary signals.

A Bayesian approach to spectral analysis of a stationary signal has been illustrated by Bretthorst [17], as well as Andrieu *et al*. [28,29]. In both cases, the observed signal was assumed to consist of the sum of unmodulated tones, such that the main parameter of interest is the frequencies of the constituent tones. Additionally, the approach adopted in both methods is to marginalise parameters which are deemed unnecessary. Parameter marginalisation refers to the process of integrating the distribution of interest over a set of nuisance parameters, or parameters which are otherwise not of interest.

As a consequence of the marginalisation process, however, the resulting posterior distribution over the frequency parameters is a multidimensional Student's t-distribution [17] and it is not possible to sample directly from this posterior distribution. The solution adopted by Andrieu *et al*. is to use Monte Carlo methods to generate samples from the posterior distribution and arrive at an estimate of the model parameters.

In the method proposed by Andrieu *et al*. [28, 29], the parameters of interest are defined as the

number of tones, their frequencies and amplitudes, as well as the observation noise variance. The amplitudes and noise variance parameters are marginalised from the posterior distribution resulting in a lower dimension state space. However, as mentioned above, this posterior distribution cannot be directly sampled from. In order to circumvent this, the authors use a reversible jump MCMC (RJMCMC) method to estimate the parameter state. The RJMCMC method, discussed in greater detail in the following chapters, is a sampler which allows comparison of states which possess different dimensions. In this estimation problem, states which possess a different number of tones are themselves of different dimensions, and thus, the RJMCMC sampler provides a means of comparing these states.

At each iteration, the reversible jump sampler attempts to improve, on average, the current estimate of the state parameters. This is done through a set of birth, death and update moves. The birth and death moves involve addition and deletion of a tone respectively, to and from the current state. The update move merely updates the values of the current model parameters through a set of conditional distributions. The reversible jump move then accepts the new proposed state according to an *acceptance probability*.

By associating the acceptance of a new state with a certain probability, the algorithm ignores new proposals which are associated with low likelihoods which get rejected due to a low probability of acceptance. At the same time, the algorithm is still capable of accepting states which are less likely, thus reducing the risk of local minima/maxima traps. By iterating these moves until convergence is achieved, the sampler is able to estimate the parameter state.

An overview of the reversible jump sampler is given in Section 3.3.2. A detailed explanation of the method and its applications are available in [29–36] as well as Sections 5.3, 6.1.1 and 6.1.2 where it is applied to estimate the number of tones and harmonics in a signal.

An extension of the method developed by Andrieu *et al*. to general frequency modulated signals was carried out by Copsey *et al*. [33]. In their method, the signal y_t, at time t, is modelled as the sum of k frequency modulated tones:

$$y_t = \sum_{j=1}^{k} a^{(j)} \cos(\omega^{(j)} t + \phi^{(j)} + m(t, \mathbf{b})) + e_t \tag{2.44}$$

where $a^{(j)}$ and $\omega^{(j)}$ denote the amplitude and frequency of the j^{th} tone, $\phi^{(j)}$ is the initial phase of the j^{th} tone, $\mathbf{b}$ is the vector of parameters describing a known frequency modulation law

$m(t, \mathbf{b})$, and the observation noise is white Gaussian, $e_t \sim \mathcal{N}(0, \sigma^2_{(e)})$ with variance $\sigma^2_{(e)}$. Copsey *et al*. use the specific example of sinusoidal frequency modulation with known modulation amplitude, but unknown modulation frequency. The authors note, however, that it is straightforward to extend the algorithm to other modulation schemes, provided a parameterised function is available for the modulation law.

As in the case of the method described by Andrieu *et al*., uninformative priors are assumed for the model parameters and the use of conjugate priors permits marginalisation of the amplitude and phase terms as well as the observation noise variance $\sigma^2_{(e)}$ from the posterior distribution. When a conjugate prior is used, the posterior distribution belongs to the same family as the prior distribution [37]. It is possible to take advantage of this when performing marginalisation by selecting prior distributions which may be integrated without difficulty.

The algorithm developed by Andrieu *et al*. then consists of using the reversible jump sampler to sample from the different dimension spaces, the different dimensions corresponding to different *model orders* or number of tones in the signal.

The application of Bayesian methods has been taken further in the analysis of polyphonic tones by Davy *et al*. [38]. The authors use a model based on Gabor atoms to represent the signal. The model accounts for a variable number of tones and harmonics, as well as allowing for inharmonicity, where the harmonics deviate from exact integer multiples of the fundamental frequency. The analysis, however, relies on segmentation of the data, such that the pitch of the signal does not change in each segment; each segment could, for example, correspond to individual chords in a piece of music.

Davy *et al*. , similar to the method described by Andrieu *et al*. [29], utilise a RJMCMC framework to determine both the number of tones as well as harmonics present in the signal. In addition to the normal birth and death moves, they also introduce multiply and divide moves which respectively double and halve the frequency of the fundamental tone. These moves are necessary to deal with ambiguities which may otherwise arise in the estimation procedure and are described in detail in Section 5.3.

The methods discussed here assume that knowledge of the signal structure is available *a priori*. For example, Copsey *et al*. assume that the frequency modulation law is known beforehand and the amplitudes of the tones are constant throughout. Such constraints limit the application of these methods to cases where the signal structure is known. In contrast, the methods cannot be

used to study the time-frequency characteristics of new and unknown signals. For such signals, sequential estimation techniques may fare better since the implicit assumption is that the signal is non-stationary and the aim is to track the time-varying frequency or spectrum of the signal. The following section looks at several online methods for time-varying frequency estimation.

2.3.2 Sequential Spectral Estimation

Online or sequential spectral estimation is necessary when dealing with non-stationary signals. The time-frequency structure of these signals may be unknown, for example as in speech or music signals where the sequence of words or notes is not known *a priori*.

One method of estimating the spectrum of a signal is to model the signal as an autoregressive (AR) process [39]. In an AR model, the observed sample $y(t)$, at discrete time t, is modelled as a linear sum of the previous samples:

$$y_t = \sum_{k=1}^{P} a^{(k)} y_{t-k} + e_t \tag{2.45}$$

where $\mathbf{a} = [a^{(1)}, \ldots, a^{(P)}]^T$ represents the AR coefficients, and e_t is zero-mean white Gaussian observation noise with variance $\sigma^2_{(e)}$. In the above formulation, the signal is assumed stationary. The aim is to determine the AR coefficients, which define the spectral density of the signal and this can be achieved, for example, by solving the Yule-Walker equations.

The AR model is often used in modelling speech signals, which are inherently non-stationary. The above AR model can be used for non-stationary signal analysis by segmenting the signal, such that each segment is approximately stationary. This is the approach followed in [40, 41] where an all-pole model of fixed order is used to approximate the spectrum of sequential segments of a speech signal. Each segment is treated independently, and the AR coefficients of each frame can be computed without considering the other frames.

A more elegant solution, however, is to specify a time-varying model for the AR coefficients. Such an approach is adopted, for example, by Schnell and Lacroix in [42]. The authors segment the signal into P non-overlapping frames, but ensure continuity rather than independence across subsequent frames. A time-varying autoregressive (TVAR) model is used such that the AR coefficients are time-varying both across frames, as well as within each frame. For the k^{th}

frame, the observed signal at time t is modelled using linear prediction as:

$$\hat{y}_{k,t} = \sum_{i=1}^{N} a_{k,t}^{(i)} y_{k,t-t_i} + e_t \tag{2.46}$$

where $\hat{y}^{k,t}$ is the estimated or predicted value of the observed sample $y^{k,t}$, $a_{k,t}^{(i)}$ represents the (time-varying) i^{th} AR coefficient for the k^{th} frame, t_i indicates the start of the k^{th} frame, and e_t is the prediction error. For the case where $t_i \geq t$, the authors specify the model to use past samples from the previous frames. Additionally, to ensure continuous functions, or smooth trajectories, of the AR coefficients, the coefficients are modelled using a linear sum of time-varying basis functions as:

$$a_{k,t}^{(i)} = \sum_{j=0}^{M} d_k^{(i,j)} \phi_{k,t}^{(j)} \tag{2.47}$$

where the time-varying basis functions are denoted by $\phi_{k,t}^{(j)}$, their weights are given by $d_k^{(i,j)}$, and additionally, the first basis function is constrained to $\phi_{k,t}^{(0)} = 1$. The authors also specify continuity across frames by specifying the constraint:

$$a_{k+1,1}^{(i)} = a_{k,N}^{(i)} \tag{2.48}$$

such that the first coefficient of the $(k+1)^{th}$ frame is identical to the last coefficient of the k^{th} frame. The authors utilise this constraint to ensure that the coefficients evolve continuously, both in time, as well as across frames. For the time-varying basis functions, Schnell and Lacroix describe linear, polynomial and trigonometric basis, with the trigonometric functions being used to model the "periodic behaviour of the glottis termination of the vocal tract." The authors employ a least mean squares (LMS) approach to minimise the prediction error and determine the TVAR coefficients.

An alternative to tracking an all-pole spectrum is possible by considering a harmonic model, where the signal is known to consist of well-defined frequencies. Such an approach is adopted, for example, by La Scala and Bitmead in tracking a single time-varying frequency in noise using the extended Kalman filter (EKF) in [43] and [44]. The authors use a non-linear model to describe noisy quadrature observations. The EKF is then used to recursively track a slowly time-varying frequency. The filter is parameterised by three tuning parameters which must be carefully chosen to achieve filter convergence and noise rejection. More recently, Hajimolahoseini *et al*. [45] have developed an EKF frequency tracker which is parameterised using a single

tuning parameter.

Frequency tracking has also been performed by tracking ridges in a time-frequency representation [46–48]. In a time-frequency representation, obtained for example using the STFT or wavelets, peaks in the distribution provide an indication of the frequencies present in the signal. If the signal is composed of a set of narrowband tones, then tracking the ridges in the time-frequency plane provides an estimate of the instantaneous frequencies of the constituent components of the signal. This is the approach followed by Streit and Barrett in [46] and Dubois *et al*. in [48] and [49] (discussed in Chapter 3) where a particle filter is used to track the time-varying peaks in the STFT.

The method developed by Streit and Barrett involves using a hidden Markov model (HMM) to track a time-varying number of narrowband components. The method relies on sequential examination of the short-time spectrum, via the STFT, to obtain frequency track estimates. The presence of a component is determined by examining the power in the bins in the Fourier transform. When certain threshold conditions are exceeded, then a tone is said to have been detected at that frequency. Using this set of detected frequencies, the Viterbi algorithm [50] is used to arrive at the maximum likelihood frequency track and the output of the filter is a set of discrete frequency tracks.

A method proposed by Clark *et al*. [51, 52] adapts a radar multi-target tracking algorithm to track a time-varying number of frequency modulated tones. In the normal target tracking framework, the aim is to assign a set of measurements to a set of target tracks. In applying the algorithm to sinusoidal tracking, the frequencies of the multicomponent signal are treated as a set of noisy measurements and the role of the target tracking algorithm is to segment these measurements into a set of frequency tracks.

A set of damped sinusoids is used to represent a small frame of audio data and the frequencies and amplitudes of the damped sinusoids which characterise the audio segment need to be estimated first. The authors utilise a subspace technique to obtain an estimate of these parameters. The frequency and amplitude estimates are then used as measurements to the probability hypothesis density (PHD) [53] filter which provides track estimates.

When performing multi-target tracking, the joint multi-target likelihood grows at an exponential rate and becomes intractable. The PHD filter offers a suboptimal solution to the multi-target tracking problem by recursively tracking the first-order moments of the multi-target posterior

distribution and peaks of the PHD are used to arrive at a set of target estimates.

Under linear and Gaussian conditions, a closed form solution to the PHD filter can be derived referred to as the Gaussian mixture PHD (GMPHD) filter [54, 55]. The GMPHD works by approximating the posterior distribution using a weighted mixture of Gaussians at time $t-1$. Then, using a linear update with Gaussian process noise, the posterior at time t is also a weighted Gaussian mixture and the mean and covariance of the distribution can be estimated sequentially using the Kalman filter. Taking the Gaussian components with the largest weights provides an estimate of the target states. Using this method Clark *et al*. apply the target tracking algorithm to arrive at track estimates. The method can thus be understood as a target tracking algorithm applied to a set of frequency estimates to arrive at a set of frequency tracks.

The methods described in this section are useful in estimating and tracking the instantaneous frequency of signals when the *global* signal structure is not known beforehand, *i.e.*, the frequency modulation law is not known *a priori*. In some estimation problems, however, the call structure may be known, for example, some bat calls can be fit to particular chirp types. In such cases, the estimation problem reduces to estimation of the chirp type and the chirp parameters. The following section reviews methods for chirp parameter estimation. These models can be used to estimate the parameters of bat calls as is discussed later in Section 3.4.2.

2.3.3 Chirp Parameter Estimation

Chirp parameter estimation can be performed using batch offline as well as sequential methods. In the case of sequential methods, the aim is to recursively refine the estimates of the set of static chirp parameters. Regardless of the approach adopted, the methods assume that the signal can be described by one of a predefined set of chirp functions. For example, a chirp could be modelled as a linear chirp with a start frequency and chirp rate, and as more data is made available sequentially, the estimates of the frequency and chirp rate are refined.

It is possible to use a frequency tracking algorithm to obtain an estimate of the IF and then estimate the chirp parameters from the instantaneous track frequencies. However, most methods discussed here attempt to estimate the chirp parameters directly from observed time-domain signal.

Linear chirp signals find widespread application in areas such as radar, sonar and communications, and consequently, they are the most commonly examined type of chirp. The observed

signal y_t, at time $t \in \{0, \ldots, T\}$, can be expressed in complex form as the sum of P chirps as:

$$y_t = \sum_{i=1}^{P} a_t^{(i)} \exp\left\{ j \left[2\pi \left(\frac{\alpha^{(i)}}{2} t^2 + f^{(i)} t \right) + \phi^{(i)} \right] \right\} + e_t \tag{2.49}$$

where $\mathbf{a}^{(i)} = [a_1^{(i)}, \ldots, a_T^{(i)}]$ is the (time-varying) amplitude of the i^{th} chirp, $\{\alpha^{(i)}, f^{(i)}, \phi_i\}$ are the frequency rate, initial frequency and phase respectively, and e_t is zero-mean, complex white Gaussian observation noise with variance $\sigma_{(e)}^2$. While the above expression describes a linear chirp, it is possible to extend the model so as to deal with higher order chirps as is discussed in [56] and Section 4.3.2. The aim is to estimate the parameters of the model $\{P, \mathbf{a}_{1:P}, \alpha_{1:P}, f_{1:P}, \phi_{0:P}\}$. This estimation problem, which has been addressed under different constraints by various authors [56–59], will be reviewed below.

Djurić and Kay [56] address parameter estimation of chirp signals with the constraint $P = 1$. While the method is described for a linear chirp, the authors note that the method is easily extended to estimate the parameters of a polynomial chirp of any order. Djurić and Kay's method involves application of a least squares method to the unwrapped phase of the signal of interest. The amplitude term is treated as a nuisance parameter and ignored from the estimation procedure. A result derived by Tretter in [60] shows that at high SNR, the observation equation can be approximated as:

$$\begin{aligned} y_t &= a \exp\left\{ j \left[2\pi \left(\frac{\alpha}{2} t^2 + f t \right) + \phi \right] \right\} + e_t & (2.50) \\ &\backsimeq a \exp\left\{ j \left[2\pi \left(\frac{\alpha}{2} t^2 + f t \right) + \phi + w_t \right] \right\} & (2.51) \end{aligned}$$

where w_t represents phase noise. Additionally, when e_t is complex Gaussian noise, then w_t is real white Gaussian noise. The frequency and phase of the signal can be estimated jointly by linear regression on the observation sequence. Djurić and Kay [56] apply this method to estimate the parameters of the linear chirp signal. They note however, that the phase must be correctly unwrapped before these estimates can be obtained. Additionally, the applicability of this method is restricted to cases with high SNR.

While the above method focuses on the case of a single chirp signal in noise, alternative methods have been proposed for parameter estimation of multiple superimposed chirp signals. In the approach followed by Liang and Arun in [57], the instantaneous frequencies and amplitudes of each chirp are first estimated sequentially using rank reduction techniques. These frequency

estimates are then used to obtain an initial estimate of the chirp parameters via a least squares fit of the chirp model. The chirp parameters are then used as initial estimates for an iterative Newton algorithm to maximise the likelihood function and arrive at a final estimate of the chirp parameters. The authors note that the initial parameter estimates are important for initialisation of the iterative Newton algorithm due to the presence of multiple maxima and minima in the likelihood function and the fact that the algorithm is locally convergent.

Saha and Kay [58] suggest a non-iterative maximum likelihood estimator for estimation of the parameters of superimposed chirps. The authors extend their algorithm applied to the estimation of superimposed tones in [61] to the problem of chirp parameter estimation. The estimation of the parameters is decoupled into a linear parameter vector composed of the complex amplitudes, and a non-linear parameter vector which consists of the chirp parameters. The non-linear parameters are estimated first, and using these estimates, the amplitudes are subsequently estimated.

When the parameters can be separated into linear and non-linear spaces, under certain conditions, it is possible to marginalise the linear parameters, as illustrated in, for example, [17], [62], as well as in Section 4.4.4. Saha and Kay refer to the marginalised likelihood as the *compressed likelihood*, which depends only on the frequency and chirp rate parameters, but not the amplitudes or observation noise variance. Maximisation of this compressed likelihood involves a multidimensional search over the parameter vector.

The authors show that a closed form expression for the frequency and chirp rates can be obtained, however, these expressions involve the evaluation of a multidimensional integral. Consequently, the authors rely on importance sampling to arrive at a Monte Carlo estimate of the chirp parameters. Importance sampling provides a means of approximating a distribution by drawing samples from some simpler distribution and weighting those samples appropriately. Using the set of weighted samples which approximate the distribution of interest, parameter estimates can be obtained by calculating expectations on the approximated distribution. Importance sampling is discussed in detail in Section 3.1 where it forms the basis of particle filters.

The marginal distributions on the parameter vectors are obtained by evaluating a joint distribution on the chirp parameters at discrete points on a (two dimensional) grid and this is used to arrive at estimates for the chirp parameters. The authors show that the method is capable of resolving closely separated signals at moderate and low SNR. Since the method relies on a

global optimisation technique, it is not prone to local minima/maxima traps as in the case of the method developed by Liang and Arun [57].

Besson *et al*. address parameter estimation for random amplitude chirp signals in [59]. The authors summarise several methods which have been developed for chirp parameter estimation for the case of constant amplitude chirp signals, as well as exponential signals with a time-varying amplitude. Their method, however, differs from these in that it is applied to the combination of the two cases, *i.e.*, chirp signals with a time-varying envelope. The authors propose a solution based on non-linear least squares (NLS) estimation which involves estimation of the parameters (ϕ, f, α), as well as the time-varying amplitudes $\mathbf{a}$. In the formulation used, the number of components present in the signal is restricted to $P = 1$. The authors propose an alternative approach which is based on the use of the higher order ambiguity function (HAF). The HAF based estimator sequentially reduces the order of the polynomial phase function of the signal, such that at each iteration, the aim is to estimate an exponential signal with a random, time-varying amplitude, and an NLS estimator is used for this purpose.

The methods described in this section provide various means for estimating chirp parameters. The following section provides an introduction to the structure of bat echolocation calls and also describes chirp types which can be used to model certain bat calls. In these instances, chirp parameter estimation can be used to estimate the call parameters, providing another option for analysis of the calls.

2.4 Overview of Bat Echolocation Signals

The aim of this section is to describe the structure and characteristics of bat echolocation calls. In order to arrive at a suitable analysis method, it is necessary for the algorithm to account for the variation which is present in the calls. Additionally, call structure varies both within as well as across species, such that any method for call analysis must refrain from making assumptions which may apply to only a few species. For example, while most calls monotonically decrease in frequency, some calls also contain *up-chirps* where the frequency increases either at the start or the end of a call. To achieve a method for accurate analysis, these deviations need to be taken into account.

The interest in bat echolocation arises from their ability to detect and differentiate between extremely small targets using biosonar. Nearly blind, bats use biosonar during flight for navi-

gation, as well as to perform target detection and classification while hunting for prey.

Aside from a highly developed auditory system, bats have evolved a complex set of calls, showing variation both within and across species [5, 63, 64]. Call variation within a species is seen when bats are in different phases of searching/tracking prey, where the structure of the call is adapted with distance to the target [5]. An example is considered here where a recording of a sequence of calls from *Pipistrellus pipistrellus* (Schreber, 1774)[2] is obtained as shown in Figure 2.1. The figure illustrates how the inter-pulse duration decreases with a simultaneous increase in call bandwidth as the bat approaches the target, with the pulse repetition frequency (PRF) being extremely high in the final stages of prey capture. Along with these changes, there is also a decrease in call duration which is not visible in the figure shown. These variations in call structure have been illustrated in greater detail by Simmons *et al*. [5] for different species.

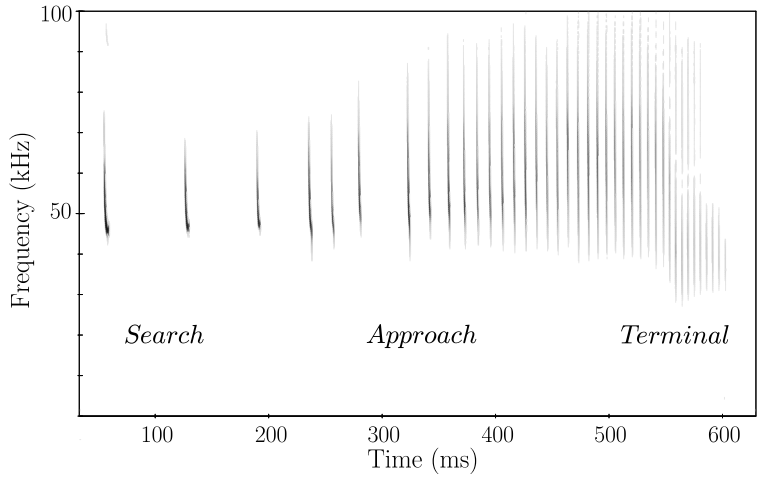


Figure 2.1: *Illustration of variation in call structure over different stages of prey pursuit for a pipistrelle bat* P. pipistrellus*. This figure is reproduced with permission from Dean Waters, University of Leeds*

Calls may be regarded as being constructed using different functional components (or combinations thereof) where the components are broadly classified as a) constant frequency (CF), b) narrowband, shallow frequency modulated (FM) (or quasi CF), and c) wideband, steep FM elements [65]. The analysis methods used for the signals thus need to be sufficiently flexible to deal with the large variability in the signal structure. Signals containing only the FM component are usually of short duration, from 0.5 to 10 ms. CF signals can be of either short duration

[2]The system for naming species by the binomial nomenclature involves specifying the latin (or scientific) name followed by the scientific authority (the person who first published the classification) in parentheses.

(1 to 10 ms) or long (10 to 100 ms) [63]. Figures 2.2-2.4 show the spectrograms of three different species of bats, with the call from *Myotis nattereri* (Kuhl, 1817) following a steep FM characteristic, the call from *P. pipistrellus* exhibiting a shallow FM structure, and the call from *Rhinolophus hipposideros* (Bechstein, 1800) being of CF type. The duration of these calls also varies significantly, with the shortest of approximately 2 ms, and the longest exceeding 50 ms. It is generally not possible to classify all calls from a single species as belonging to a single type, *i.e.*, shallow FM, steep FM, or CF, since the call structure adapts according to the situation and the boundaries between these categories can be ambiguous.

It has been observed that the structure of bat calls is substantially influenced by ecological conditions (habitat type, foraging mode, diet) [64], leading to inter-species variation of calls. In uncluttered or open spaces, sonar echoes from surrounding foliage (or other clutter) are either undetectable or sufficiently far apart from echoes of potential prey. To take advantage of this situation, bats use intense, narrowband, shallow FM calls of long duration which permits long-range detection of prey.

Bats foraging in background-cluttered space must deal with echoes from clutter which closely follow those of prey. They use calls having a short-duration wideband component followed by a longer narrowband FM component which allows a compromise between medium-range detection of prey and target localisation and classification.

Bats which forage in highly-cluttered spaces must be able to detect prey-echoes that are partially or completely buried in echoes of the surrounding clutter. These bats generally use long-duration CF calls ending in a downward-modulated, wideband FM tail (CF-FM calls). CF calls allow bats to decode Doppler information present in the echo return.

While calls can be broadly classified as FM or CF calls, in some cases, it is possible to fit the time-frequency representation of the call to a particular chirp type, *i.e.*, the frequency modulation law of the call can be fit to a particular chirp. Some of these chirps are examined in the following section and chirp parameter estimation of echolocation calls will be considered in the following chapter.

Chirp Structures of Echolocation Calls

A study of the various calls allows classification under several categories of chirp signals. Steep, wideband FM signals may be modelled as linear chirps. Doppler tolerant waveforms follow

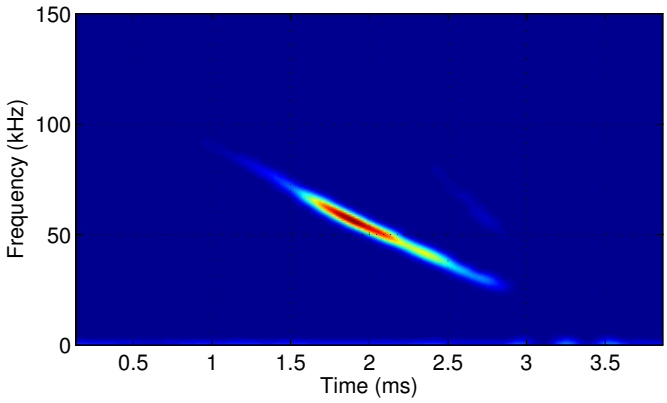


Figure 2.2: *Spectrogram of a call emitted by* M. nattereri

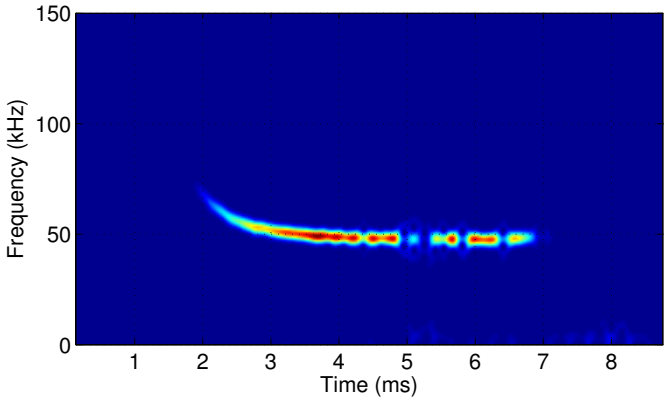


Figure 2.3: *Spectrogram of a call emitted by* P. pipistrellus

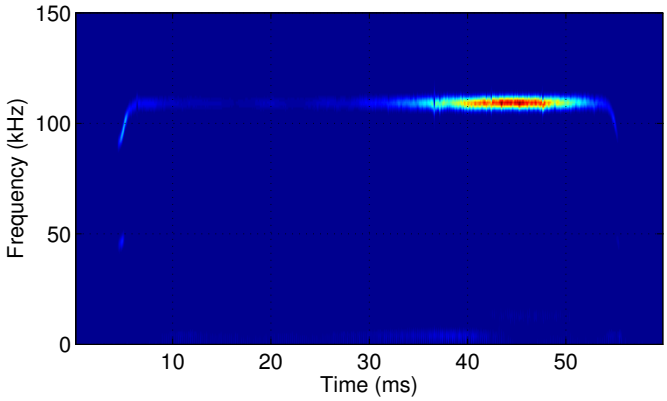


Figure 2.4: *Spectrogram of a call emitted by* R. hipposideros

linear-period modulation; bat chirps of the little brown bat, *Myotis lucifugus* (LeConte, 1831), have been found to have similar traits, but follow a logarithmic-time modulation [66, 67]. Exponential chirps [67] have been used to model bat calls for the big brown bat, *Eptesicus fuscus* (Palisot de Beauvois, 1796) [68]. For an exponential chirp, critical parameters were identified as starting and terminal frequency of the first harmonic, amplitude envelopes of the first three harmonics, and decay constant of the exponentially falling frequency sweep.

If the frequency at time instant t is given by f_t, then the chirp is given by

$$z_t = a_t \cdot \cos(2\pi f_t t + \phi) \tag{2.52}$$

where $z_t$is the observation, ϕ is the initial phase, f_t is the time-varying frequency and a_t is the amplitude at time t. The term a_t can be used to describe an amplitude modulation on the waveform. Echoes from targets may contain significant amplitude and/or frequency modulation terms.

For a linear chirp, the evolution of the instantaneous frequency is given as:

$$f_t = f_0 + \frac{m}{2} \cdot t \tag{2.53}$$

where f_0 is the initial frequency and m is the chirp rate.

Frequency of an exponentially modulated chirp is given by the equation [67]:

$$f_t = B \exp\left(\frac{t}{T}\right) + f_a \tag{2.54}$$

where T is the time decay constant (negative), f_a is the asymptotic frequency approached when t is large and B is the bandwidth (starting frequency minus f_a).

Logarithmic-time modulation is described as [67]:

$$f_t = -F \ln\left[\frac{t - t_a}{D}\right] + f_e \tag{2.55}$$

where f_e is the terminal frequency in the call, F is the frequency decay constant (negative), t_a is the time asymptote (< 0, *i.e.*, before the emission actually starts at $t = 0$) and D is the sum of call duration and the absolute value of t_a.

In the following chapters, parameter estimation of bat echolocation calls will be considered, as

well as the limitations of the models and methods considered for the analysis of echolocation calls.

2.5 Summary

The aim of this chapter has been to provide a background to the topic of time-frequency analysis. Central to time-frequency analysis is the concept of instantaneous frequency (IF), which is defined as the derivative of the phase. The analytic signal provides a way of representing a real signal in terms of its complex counterpart while ensuring that the spectrum is non-negative. However, as noted by Cohen [7], there are several issues which appear to be at odds with this definition.

It is possible to determine the time-varying spectrum of a signal by examination of its time-frequency distribution. The STFT, wavelet methods and the Wigner distribution are considered in some detail in Section 2.2. These methods constitute a non-parametric approach to the time-frequency problem by estimating the spectrum over all possible frequencies at different instants of time.

Parametric time-frequency approaches, by contrast, can provide a concise set of discrete frequencies that are present at different instants of time. These methods, however, rely on an ability to characterise the signal using a suitable model. To this end, TVAR models and harmonic models have found widespread application; sequential Monte Carlo (SMC) methods relying on harmonic models will be considered further in the rest of this thesis.

Another aspect considered in this chapter is chirp parameter estimation. This can be considered as a form of parametric time-frequency analysis, where the chirp type and chirp parameters define the time-varying frequency of the signal. As parametric methods, they rely on a suitable chirp model being available for the signal of interest, the signals of interest in the present case being bat echolocation calls.

In the analysis of bat echolocation calls considered here, it is the time-frequency structure of the call that is of interest. The methods considered here suffer from certain limitations. The use of non-parametric approaches requires post-processing of the time-frequency distribution to obtain the time-varying frequencies present in the signal. Alternatively, a parametric approach requires that a comprehensive model be available which can describe the time-frequency be-

haviour of calls from different species. Section 2.4 lists various chirp models which have been fit to echolocation calls. The models, however, do not consider the presence of higher harmonics or amplitude modulation of the call. The model fitting problem is further complicated by the fact that there are over 800 species in existence [69] and there is a significant amount of inter- as well as intra-species call variability.

The difficulties encountered with chirp parameter estimation can be overcome to some extent by first sequentially estimating the frequencies present in the call and then performing chirp parameter estimation on those frequency estimates. The following chapter examines existing SMC methods for time-varying frequency analysis. Chapters 4 and 5 seek to address the limitations of current methods by incorporating sufficient flexibility in the signal model and developing a new and robust technique for sequential frequency estimation.

Chapter 3
Sequential Monte Carlo Methods for Time-Frequency Analysis

Introduction

The previous chapter introduced the basic concepts of time-frequency analysis as well as examining several non-parametric and parametric approaches to the time-frequency estimation problem. This chapter explores a subclass of parametric methods which rely on sequential Monte Carlo (SMC) methods.

The following section first introduces the basics of particle filtering or SMC methods. This is followed by a discussion of the Rao-Blackwellised particle filter (RBPF) which facilitates marginalisation of parameters within the particle filter framework. Marginalisation is possible when the parameter space can be separated into a set of linear and non-linear parameters such as when the observation equation is linear in the amplitude term. In such cases, the particle filter is used to estimate the non-linear parameters while a Kalman filter can be used to estimate the linear parameters.

Prior to the examination of methods for sequential time-frequency analysis, jump Markov system (JMS) and reversible jump MCMC (RJMCMC) methods are introduced. These methods offer a framework for estimation of an unknown *model order*. This unknown model order, in the context of time-varying frequency analysis, could represent the time-varying number of tones and/or harmonics in the signal, or a time-varying number of AR coeffiecients.

Section 3.4 looks at several SMC methods for tracking a time-varying (and unknown) number of tones. In these methods, the underlying signal model is either a TVAR model or a harmonic model. Specifically, Section 3.5 examines, in detail, a RBPF method developed by Dubois and Davy for time-varying frequency estimation [70]. The performance of this method is subsequently compared with a new proposed method which is developed in Chapter 5.

3.1 Introduction to Particle Filtering

The Kalman filter [71, 72] provides the optimal solution to linear dynamic state estimation problems. The Kalman filter estimates the unknown state sequentially by tracking the mean and variance of the posterior Gaussian density.

The extended Kalman filter (EKF) [72] and the unscented Kalman filter (UKF) [73, 74] are modifications to the basic Kalman filtering algorithm which then allow the algorithm to deal with non-linearity in the model. The use of a non-linear state update and/or observation model results in the posterior distributions being non-Gaussian and the Kalman filter cannot be used as is.

The EKF operates by approximating the non-linear model locally using a linear model. The linearisation process involves evaluation of Jacobians which replace the state update and observation matrices used in the Kalman filter. The EKF has been shown to perform well in the presence of mild non-linearity in the model, but performs poorly in the presence of high degrees of non-linearity [72]. As discussed in the previous chapter, the EKF has been used to estimate the time-varying frequency of a sinusoid in noise.

Unlike the EKF, the UKF is capable of dealing with non-linearity to a greater extent. The UKF propagates a minimal and deterministic set of points through the non-linearity which results in preservation of the first two moments, *i.e.*, the mean and covariance. Additionally, the need to compute the Jacobian as in the case of the EKF is removed since the linearisation process is avoided. One significant drawback of the (extended/unscented) Kalman filters is that they are entirely unsuitable for the estimation of multi-modal distributions as noted by Ristic *et al*. [72]. Since these filters characterise the distribution of interest using a mean and covariance, they are unable to deal with multimodal distributions.

The limitations in dealing with non-linearity and multimodal distributions have motivated the use of SMC methods. The following sections provide an overview of particle filter methods and their application to frequency estimation will be subsequently considered.

3.1.1 Monte Carlo Integration

Particle filters [72,75–78] are a class of suboptimal filters that provide a method for "implementing a recursive Bayesian filter by [using] Monte Carlo simulations" [72]. While the Kalman fil-

ter and its derivatives merely track the moments of the Gaussian posterior distribution, particle filters track the posterior distribution itself. This is done by evaluating the distribution using a set of discrete samples, or particles, over the parameter space. This sampling is performed on a non-uniform and adaptive grid such that it becomes possible to sample the distribution more finely in regions of interest, offering significant benefits over a uniform grid size. Particle filters are thus capable of dealing with non-linear functions, non-Gaussian noise, as well as multi-modal posterior distributions, making them a flexible and attractive option for such problems.

The filter relies on a state space model to decide which region of the parameter space needs to sampled. The state space equations are written as [72]:

$$\boldsymbol{\psi}_k = g_{k-1}(\boldsymbol{\psi}_{k-1}) + \mathbf{v}_{k-1} \tag{3.1a}$$

$$\mathbf{x}_k = h_k(\boldsymbol{\psi}_k) + \mathbf{w}_k \tag{3.1b}$$

where $\boldsymbol{\psi}_k$ represents the (vector) state that is required to be estimated, $\mathbf{x}_k$ is the noisy observations, $g_{k-1}(\cdot)$ is a possibly non-linear function describing the state update, $h_k(\cdot)$ is a possibly non-linear observation function, $\mathbf{v}_{k-1}$ and $\mathbf{w}_k$ represent the process and observation noise terms respectively.

Let $\boldsymbol{\Psi}_k = \boldsymbol{\psi}_{0:k} = \{\boldsymbol{\psi}_0, \ldots, \boldsymbol{\psi}_k\}$ denote the sequence of states up to time k, and $\mathbf{X}_k = \mathbf{x}_{1:k} = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$, the sequence of observations. The aim of the filter is to recursively estimate the posterior distribution $p(\boldsymbol{\Psi}_k \mid \mathbf{X}_k)$. Of further interest may be the marginal or filtering distribution $p(\boldsymbol{\psi}_k \mid \mathbf{X}_k)$, as well as the expectation [72, 79]:

$$I = \int f(\boldsymbol{\Psi}_k) p(\boldsymbol{\psi}_k \mid \mathbf{X}_k) d\boldsymbol{\Psi}_k \tag{3.2}$$

where $f(\cdot)$ indicates some function of interest. Depending on the quantity of interest, this function can, for example, denote the conditional mean or conditional covariance [79].

When it is possible to draw $N \gg 1$ samples, $\{\boldsymbol{\Psi}_k^{(i)}; i = 1, \ldots, N\}$, from the distribution $p(\boldsymbol{\Psi}_k \mid \mathbf{X}_k)$, then equation (3.2) can be approximated using the Monte Carlo estimate [72]:

$$I \approx \hat{I}_N = \frac{1}{N} \sum_{i=1}^{N} f(\boldsymbol{\Psi}_k^{(i)}) \tag{3.3}$$

where the estimate $\hat{I}_N$ is unbiased if the samples $\mathbf{\Psi}_k^{(i)}$ are independent. Additionally, under the strong law of large numbers, the Monte Carlo estimate, $\hat{I}_N$, obtained using equation (3.3) will almost surely converge to I, $\hat{I}_N \rightarrow I$, as the number of samples approaches infinity, $N \rightarrow \infty$ [72, 79]. Ristic *et al*. [72] also note that the rate of convergence is independent of the dimension of the integrand, in contrast with deterministic numerical integration methods where the rate of convergence decreases as the dimension increases. However, it should be noted that the number of samples needed to achieve the desired accuracy of the estimate will still increase with the dimension of the parameter space. This is illustrated in Chapter 7 in the context of frequency estimation of a FM signal.

Using equation (3.3), it becomes possible to compute the expectation of the function $f(\mathbf{\Psi}_k)$, thus allowing evaluation of, for example, the minimum mean squared error (MMSE) estimate. In practice, however, it can be difficult to sample directly from $p(\mathbf{\Psi}_k \mid \mathbf{X}_k)$. In order to draw samples from the posterior, importance sampling [36, 72, 79] is utilised.

3.1.2 Importance Sampling

Importance sampling works by drawing samples from a simpler importance or proposal distribution. If these samples are correctly weighted, then the Monte Carlo estimate of equation (3.3) can still be obtained. Let $q(\mathbf{\Psi} \mid \mathbf{X})$ denote the importance distribution, such that:

$$p(\mathbf{\Psi} \mid \mathbf{X}) > 0 \implies q(\mathbf{\Psi} \mid \mathbf{X}) > 0 \qquad \forall \ \mathbf{\Psi} \in \mathbb{R}^{n_{\mathbf{\Psi}}} \tag{3.4}$$

where $n_{\mathbf{\Psi}}$ denotes the dimension of the state $\mathbf{\Psi}$. This condition implies that the importance distribution must cover the support of the posterior distribution which is a necessity for importance sampling [72]. Equation (3.2) can be rewritten as:

$$I = \int f(\mathbf{\Psi}) \frac{p(\mathbf{\Psi} \mid \mathbf{X})}{q(\mathbf{\Psi} \mid \mathbf{X})} q(\mathbf{\Psi} \mid \mathbf{X}) d\mathbf{\Psi}. \tag{3.5}$$

If $N \gg 1$ samples are drawn from the importance distribution, $\{\mathbf{\Psi}^{(i)} \sim q(\mathbf{\Psi}); i = 1, \ldots, N\}$, then the above integral can be approximated as:

$$I \approx \hat{I}_N = \frac{1}{N} \sum_{i=1}^{N} f(\mathbf{\Psi}^{(i)}) \frac{p(\mathbf{\Psi}^{(i)} \mid \mathbf{X})}{q(\mathbf{\Psi}^{(i)} \mid \mathbf{X})} = \frac{1}{N} \sum_{i=1}^{N} f(\mathbf{\Psi}^{(i)}) \hat{w}^{(i)} \tag{3.6}$$

with

$$\hat{w}^{(i)} = \frac{p(\mathbf{\Psi}^{(i)} \mid \mathbf{X})}{q(\mathbf{\Psi}^{(i)} \mid \mathbf{X})} \tag{3.7}$$

where $\hat{w}^{(i)}$ denotes the importance weight associated with the i^{th} sample. In cases where the normalising constant of the posterior density $p(\mathbf{\Psi} \mid \mathbf{X})$ is unknown, the importance weights need to be normalised and the Monte Carlo integral is computed as:

$$\hat{I}_N = \frac{\frac{1}{N}\sum_{i=1}^{N} f(\mathbf{\Psi}^{(i)})\hat{w}^{(i)}}{\frac{1}{N}\sum_{j=1}^{N} \hat{w}^{(j)}} = \sum_{i=1}^{N} f(\mathbf{\Psi}^{(i)})w^{(i)} \tag{3.8}$$

where

$$w^{(i)} = \frac{\hat{w}^{(i)}}{\sum_{j=1}^{N} \hat{w}^{(j)}} \tag{3.9}$$

denotes the normalised importance weight [72].

Importance sampling can be extended to a sequential framework allowing recursive estimation of a dynamic state. The following section introduces sequential importance sampling (SIS) which forms the basis of particle filters.

3.1.3 Sequential Importance Sampling

A particle filter approximates the posterior distribution by a discrete set of weighted samples using SIS. Using importance sampling, the time-varying posterior distribution at time k is represented by the set of N normalised weighted particles, denoted as $\{\mathbf{\Psi}_k^{(i)}, w_k^{(i)}\}, i \in \{1, \ldots, N\}$. The posterior distribution $p(\mathbf{\Psi}_k \mid \mathbf{X}_k)$, can be approximated as:

$$p(\mathbf{\Psi}_k \mid \mathbf{X}_k) \approx \sum_{i=1}^{N} w_k^{(i)} \delta(\mathbf{\Psi}_k - \mathbf{\Psi}_k^{(i)}) \tag{3.10}$$

where $\delta(\cdot)$ denotes the Dirac delta function and the importance weights are evaluated as:

$$w_k^{(i)} = \frac{p(\mathbf{\Psi}_k^{(i)} \mid \mathbf{X}_k)}{q(\mathbf{\Psi}_k^{(i)} \mid \mathbf{X}_k)}. \tag{3.11}$$

The approximation of equation (3.10) may be interpretted as a means of approximating the

cumulative distribution function (CDF) rather than the probability distribution function (PDF). The true CDF, which may be a continuous function, is approximated using a set of weighted discrete delta functions at points in the state space denoted by $\mathbf{\Psi}_k^{(i)}$.

The aim is to then obtain a recursive relation such that it is possible to obtain $p(\mathbf{\Psi}_k \mid \mathbf{X}_k)$ given $p(\mathbf{\Psi}_{k-1} \mid \mathbf{X}_{k-1})$ on receipt of the new sample $\mathbf{x}_k$. This can be achieved by suitably factorising the importance and posterior distributions using the properties of causality and independence implicit in the model state space as shown respectively in equations (3.12) and (3.13)-(3.18) [72]:

$$q(\mathbf{\Psi}_k \mid \mathbf{X}_k) = q(\boldsymbol{\psi}_k \mid \mathbf{\Psi}_{k-1}, \mathbf{X}_k) q(\mathbf{\Psi}_{k-1} \mid \mathbf{X}_{k-1}) \tag{3.12}$$

such that the new state $\mathbf{\Psi}_k^{(i)}$ is achieved by augmenting the existing state $\mathbf{\Psi}_{k-1}^{(i)}$ with the new sample $\boldsymbol{\psi}_k \sim q(\boldsymbol{\psi}_k \mid \mathbf{\Psi}_{k-1}, \mathbf{X}_k)$.

Applying Bayes's rule to the posterior distribution $p(\mathbf{\Psi}_k \mid \mathbf{X}_k)$ yields:

$$p(\mathbf{\Psi}_k \mid \mathbf{X}_k) = \frac{p(\mathbf{x}_k \mid \mathbf{\Psi}_k, \mathbf{X}_{k-1}) p(\mathbf{\Psi}_k \mid \mathbf{X}_{k-1})}{p(\mathbf{x}_k \mid \mathbf{X}_{k-1})} \tag{3.13}$$

$$= \frac{p(\mathbf{x}_k \mid \mathbf{\Psi}_k, \mathbf{X}_{k-1}) p(\boldsymbol{\psi}_k \mid \mathbf{\Psi}_{k-1}, \mathbf{X}_{k-1}) p(\mathbf{\Psi}_{k-1} \mid \mathbf{X}_{k-1})}{p(\mathbf{x}_k \mid \mathbf{X}_{k-1})} \tag{3.14}$$

The above expression can be simplified by incorporating the assumptions in the model defined by the state space equations (3.1a) and (3.1b). The observation sample depends only on the current state allowing the simplification:

$$p(\mathbf{x}_k \mid \mathbf{\Psi}_k, \mathbf{X}_{k-1}) = p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) \tag{3.15}$$

Further, the state update is described using a Markov chain with dependence on only the value at the previous instant. Incorporating this allows the additional simplification:

$$p(\boldsymbol{\psi}_k \mid \mathbf{\Psi}_{k-1}, \mathbf{X}_{k-1}) = p(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_{k-1}) \tag{3.16}$$

Using the above expressions, the posterior distribution can then be recursively factorised as [72]:

$$p(\mathbf{\Psi}_k \mid \mathbf{X}_k) = \frac{p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) p(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_{k-1})}{p(\mathbf{x}_k \mid \mathbf{X}_{k-1})} p(\mathbf{\Psi}_{k-1} \mid \mathbf{X}_{k-1}) \tag{3.17}$$

$$\propto p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) p(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_{k-1}) p(\mathbf{\Psi}_{k-1} \mid \mathbf{X}_{k-1}) \tag{3.18}$$

Equation (3.11), which describes the importance weights, can be rewritten in terms of equations (3.12) and (3.18) such that the importance weight is obtained in recursive form as:

$$w_k \propto \frac{p(\mathbf{x}_k \mid \boldsymbol{\psi}_k)p(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_{k-1})p(\boldsymbol{\Psi}_{k-1} \mid \mathbf{X}_{k-1})}{q(\boldsymbol{\psi}_k \mid \boldsymbol{\Psi}_{k-1}, \mathbf{X}_k)q(\boldsymbol{\Psi}_{k-1} \mid \mathbf{X}_{k-1})} \tag{3.19}$$

$$= w_{k-1}\frac{p(\mathbf{x}_k \mid \boldsymbol{\psi}_k)p(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_{k-1})}{q(\boldsymbol{\psi}_k \mid \boldsymbol{\Psi}_{k-1}, \mathbf{X}_k)} \tag{3.20}$$

where the particle index (i) has been dropped for clarity. The filtering density is then approximated using N particles as [72]:

$$p(\boldsymbol{\psi}_k \mid \mathbf{X}_k) \approx \sum_{i=1}^{N} w_k^{(i)} \delta(\boldsymbol{\psi}_k - \boldsymbol{\psi}_k^{(i)}). \tag{3.21}$$

3.1.4 Degeneracy and Resampling

It is desirable to have the sampling distribution as similar as possible to the posterior distribution, and ideally, it would be possible to sample from the posterior itself. Factorising the importance distribution as in equation (3.12) leads to an increase in the variance of the importance weights over time [72]. After successive iterations, all except a single particle will have negligible weight, resulting in a problem referred to as *degeneracy*. This leads to a situation where a large number of particles in the filter make no contribution to the estimate of the posterior distribution.

Figure 3.1 illustrates the effect of degeneracy when a particle filter is applied to track the frequency of a signal. In the example, a basic harmonic model, described in detail in Chapter 4, is used to describe an FM signal. The particle filter is used to track the time-varying frequency of this signal. It is seen from the figure that a few iterations are sufficient to reduce the weights of most particles to zero, resulting in poor tracking performance.

A method of mitigating degeneracy is to replace particles having low weights with particles associated with large weights through a process termed *resampling*. Resampling involves drawing samples from the posterior distribution at the current time instant. This is done through a process similar to the inverse transform [36], which involves drawing samples from a distribution whose CDF is known.

The cumulative sum of weights (CSW) exists in the range $[0, 1]$. Thus, if a uniform random number is generated, $u \sim \mathcal{U}[0, 1]$, the inverse transform would involve selecting the parti-

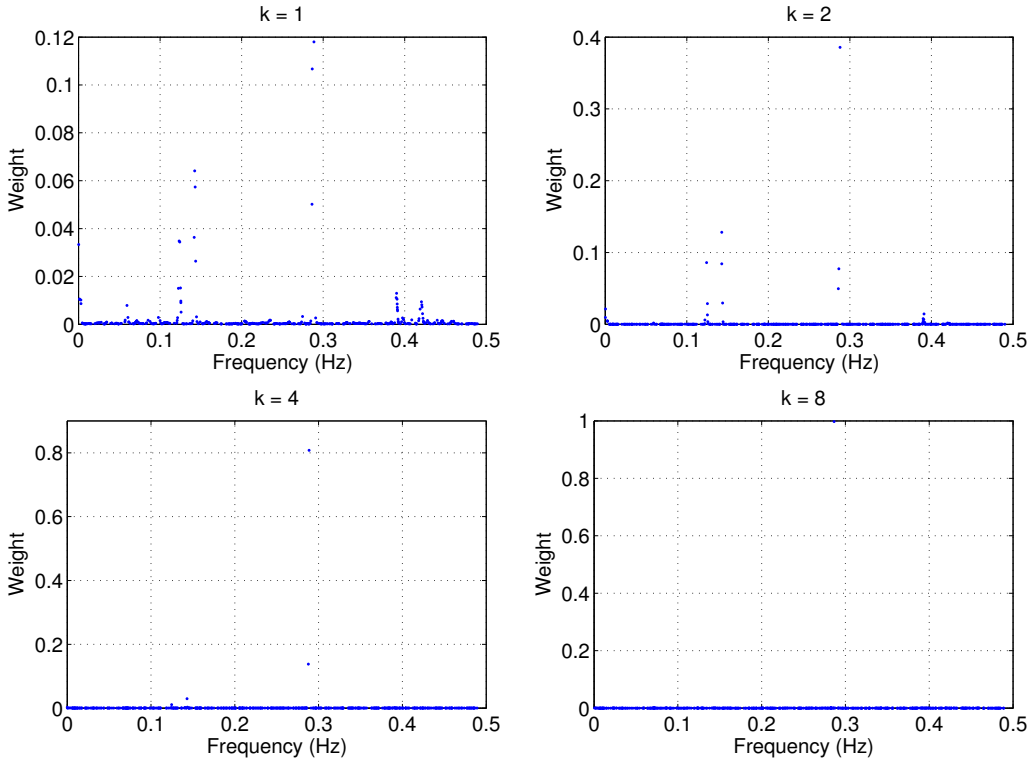


Figure 3.1: *Example of a filter suffering from degeneracy. As the time step k increases, the weights of all except a few particles decreases to zero.*

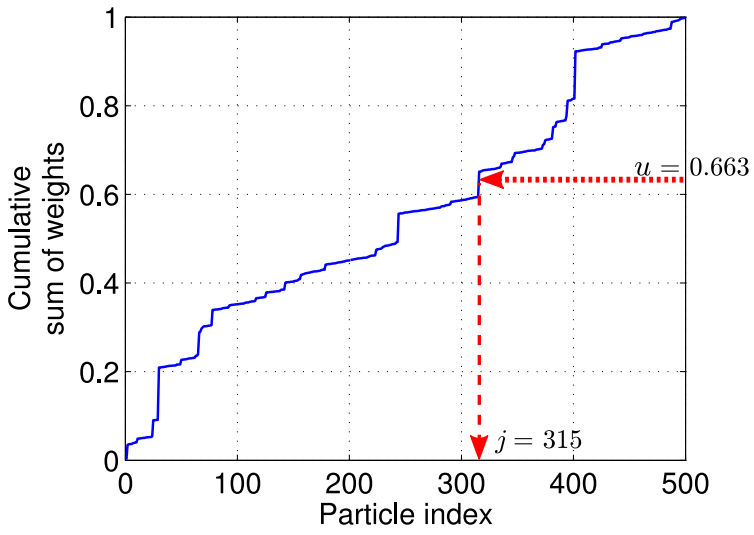


Figure 3.2: *Illustration of the process of multinomial resampling. A uniform random number u is drawn and the particle corresponding to the CSW is selected. Particles with large weights will get selected more often than particles with low weights.*

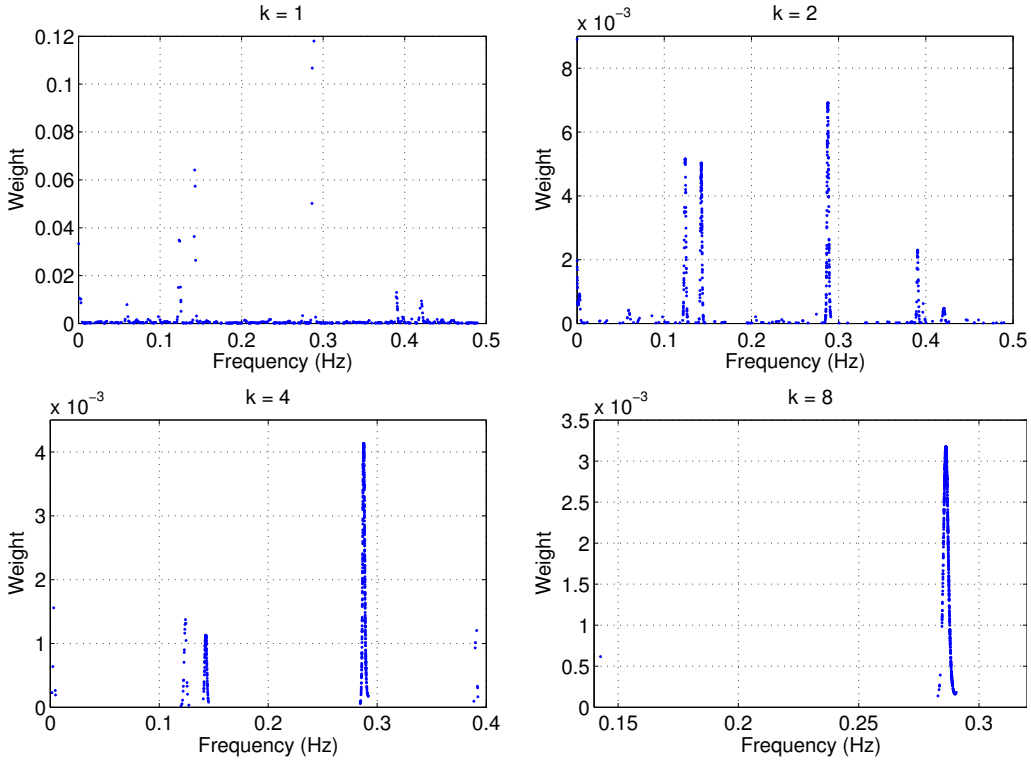


Figure 3.3: *Addition of the resampling step prevents degeneracy.*

cle index associated with that value of CSW. The CSW (also known as the cumulative mass function (CMF)) is computed from the summation of the particle weights and the resampling operation is performed by obtaining the index of the j^{th} particle which corresponds to the CSW u. By drawing N such uniformly distributed samples, N independently and identically distributed (i.i.d.) particles can be drawn from the distribution and the new weights associated with these particles will be uniform, *i.e.* $\frac{1}{N}$. This scheme for resampling is referred to as *multinomial resampling* [80, 81].

Figure 3.2 illustrates this process where a sample u is drawn and the particle index j corresponding to the CSW of u is selected. Figure 3.3 shows the filter tracking the frequency of a signal, but now with the use of the resampling step. It is seen that better characterisation of the posterior is obtained with the use of resampling.

Multinomial resampling is an expensive operation due to the inverse transform which operates on N independent uniform samples. In order to improve the efficiency of the resampling step, alternate resampling procedures can be used such as *stratified*, *systematic* or *residual* resampling [80, 81]. Of all these schemes, systematic resampling is widely used due to its ease of implementation.

In the case of systematic resampling, the N uniform samples are generated as [81]:

$$u_j = \frac{j - 1 + \tilde{u}}{N}, \quad \tilde{u} \sim \mathcal{U}[0, 1], j \in \{1, \ldots, N\}. \tag{3.22}$$

This is equivalent to partitioning the range $[0, 1]$ into N equal intervals and drawing a sample in each interval. However, the N uniform samples are not independent since they are all generated using a single random variable $\tilde{u}$ and consequently, Douc *et al*. [80], as well as Hol *et al*. [81] note that a performance analyis of the scheme is difficult. The authors in both cases also point out that the performance of systematic resampling is on par with the other resampling schemes, and thus, only the systematic resampling scheme is used in this work. The steps involved in the systematic resampling scheme are reproduced from [72] in Algorithm 1.

Although resampling solves the degeneracy problem, it reduces particle diversity since the filter contains only a few unique particles leading to sample impoverishment [72]. Sample impoverishment can be limited by *rejuvenation* of the particles through the use of a Markov chain Monte Carlo (MCMC) move step in the filter or the resample-move algorithm to subsequently increase particle diversity [72]. The MCMC move step involves applying a Metropolis-Hastings sampler

to the existing particles so as to draw particles from the posterior distribution thus increasing the number of unique particles in the filter.

Algorithm 1 Systematic Resampling algorithm.

procedure RESAMPLE($\{\boldsymbol{\psi}_k^{(i)}, w_k^{(i)}\}_{i=1}^N$)
  Evaluate CSW:
  $c_1 = w_k^{(1)}$
  for $i = 2 : N$ **do**
  $c_i = c_{i-1} + w_k^{(i)}$
  end for

  Set $i = 1$ and draw $\tilde{u} \sim \mathcal{U}\left[0, \frac{1}{N}\right]$
  for $j = 1 : N$ **do**
    Set $u_j = \tilde{u} + \frac{1}{N}(j-1)$
    while $u_j > c_i$ **do**
      Increment $i = i + 1$
    end while
    Assign $\boldsymbol{\psi}_k^{(j)\star} = x_k^{(i)}$, $w_k^{(j)} = \frac{1}{N}$
  end for

  Return $\{\boldsymbol{\psi}_k^{(i)\star}, w_k^{(i)}\}_{i=1}^N$
end procedure

3.1.5 SIR Filter

The sequential importance resampling (SIR) filter is the simplest particle filter. The SIR filter is introduced first and the following section describes a variant of this filter called the RBPF which addresses parameter marginalisation within the particle filtering framework.

The SIR filter consists of SIS combined with resampling at each iteration. Additionally, by selecting the transition prior $p(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_{k-1})$ as the importance distribution (also referred to as the sampling distribution), the weight update equation simplifies to:

$$\hat{w}_k = p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) \tag{3.23}$$

where the dependence on the previous weight is removed since resampling is performed on each iteration. Algorithm 2 lists the steps involved in the implementation of the SIR filter.

As noted in [72], due to the importance distribution being independent of the observations $\mathbf{x}_k$, the filter can be inefficient and sensitive to outliers since the state space is explored without

incorporating knowledge from the observations. The advantage of the SIR filter comes from the ease of implementation since the importance weights are readily obtained from the likelihood function $p(\mathbf{x}_k \mid \boldsymbol{\psi}_k)$ [72].

Algorithm 2 SIR particle filter.

Initialise:
$\boldsymbol{\psi}_0^{(i)} \sim p(\boldsymbol{\psi}_0) \mid_{i=1}^{N}$
1: **for** $k = 1 : k_{max}$ **do**
2: Process-update:
$\boldsymbol{\psi}_k^{(i)} \sim p(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_{k-1}^{(i)}) \mid_{i=1}^{N}$
3: Evaluate importance weight:
$\hat{w}_k^{(i)} = p(\mathbf{x}_k \mid \boldsymbol{\psi}_k^{(i)})$
4: Normalise weights:
$w_k^{(i)} = T^{-1}\hat{w}_k^{(i)} \mid_{i=1}^{N}$ where $T = \sum_{i=1}^{N} \hat{w}_k^{(i)}$
5: Resample:
$\{\boldsymbol{\psi}_k^{(i)}, w_k^{(i)}\} = \text{RESAMPLE}[\{\boldsymbol{\psi}_k'^{(i)}, w_k^{(i)}\}_{i=1}^{N}]$
6: **end for**

3.2 Parameter Marginalisation: The Rao-Blackwellised Particle Filter

Parameter marginalisation refers to the process of integrating a distribution of interest over a nuisance parameter. This process, which is sometimes referred to as Rao-Blackwellisation, results in an estimator with lower variance and is thus sometimes referred to as variance-reduction methods.

Marginalisation of parameters is possible, and desirable, in some implementations of a particle filter. For large parameter state spaces, sampling can be inefficient, and marginalisation of parameters additionally reduces the variance of the estimator [36]. Marginalisation involves taking advantage of the structure of the state space models to analytically marginalise some of the parameters. This section reviews the Rao-Blackwellised particle filter (RBPF) as applied to the class of problems which are conditionally linear and Gaussian. An alternative marginalisation process will be developed in Chapter 5 and a comparison between the methods will be illustrated there.

The RBPF [70, 76, 82–85] provides a means for parameter marginalisation when applied to conditionally linear Gaussian state space models. In such models, it is possible to partition

the state such that one set of parameters can be modelled using a linear Gaussian model conditional on the other set of parameters. Marginalisation of parameters reduces the dimension of the parameter state space, and, as mentioned earlier, leads to an estimator with lower variance. Consequently, the RBPF is an attractive option for application to conditionally Gaussian models.

Consider the state space equations:

$$\boldsymbol{\psi}'_k = g_{k-1}(\boldsymbol{\psi}'_{k-1}) + \mathbf{v}_{(\psi'),k-1} \tag{3.24a}$$

$$\boldsymbol{\xi}_k = \mathbf{A}\boldsymbol{\xi}_{k-1} + \mathbf{B}\mathbf{v}_{(\xi),k-1} \tag{3.24b}$$

$$\mathbf{x}_k = \mathbf{C}\boldsymbol{\xi}_k + \mathbf{D}\mathbf{w}_{(x),k} \tag{3.24c}$$

where $\mathbf{v}_{(\psi'),k-1}$, $\mathbf{v}_{(\xi),k-1}$ and $\mathbf{w}_{(x),k}$ denote additive noise terms and the subscript indicates that the noise applies to the states $\boldsymbol{\psi}'$, $\boldsymbol{\xi}$ and $\mathbf{x}$ respectively; the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ may be non-linear functions of the state $\boldsymbol{\psi}'_k$ [70]. In the above equations, conditional on the state $\boldsymbol{\psi}'_k$, the observation is linear in $\boldsymbol{\xi}_k$ and possesses a Gaussian distribution. Additionally, the parameter space has been partitioned as $\boldsymbol{\psi}_k = \{\boldsymbol{\psi}'_k, \boldsymbol{\xi}_k\}$ such that [82]:

$$p(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_{k-1}) = p(\boldsymbol{\xi}_k \mid \boldsymbol{\psi}'_{k-1:k}, \boldsymbol{\xi}_{k-1}) p(\boldsymbol{\psi}'_k \mid \boldsymbol{\psi}'_{k-1}) \tag{3.25}$$

and, conditional on $\boldsymbol{\Psi}'_k$, $p(\boldsymbol{\Xi}_k \mid \mathbf{X}_k, \boldsymbol{\Psi}'_k)$ can be evaluated analytically, where $\boldsymbol{\Psi}'_k = \boldsymbol{\psi}'_{0:k}$ and $\boldsymbol{\Xi}_k = \boldsymbol{\xi}_{0:k}$. If $\mathbf{v}_{(\xi),k-1}$ and $\mathbf{w}_{(x),k}$ are zero-mean Gaussian, as is assumed here, then $p(\boldsymbol{\Xi}_k \mid \mathbf{X}_k, \boldsymbol{\Psi}'_k)$ is conditionally Gaussian.

The posterior distribution can be written as [82]:

$$p(\boldsymbol{\Psi}'_k, \boldsymbol{\Xi}_k \mid \mathbf{X}_k) = p(\boldsymbol{\Xi}_k \mid \mathbf{X}_k, \boldsymbol{\Psi}'_k) p(\boldsymbol{\Psi}'_k \mid \mathbf{X}_k) \tag{3.26}$$

where $p(\boldsymbol{\Psi}'_k \mid \mathbf{X}_k)$ can be further expanded as [82]:

$$p(\boldsymbol{\Psi}'_k \mid \mathbf{X}_k) = \frac{p(\mathbf{x}_k \mid \boldsymbol{\Psi}'_k, \mathbf{X}_{k-1}) p(\boldsymbol{\psi}'_k \mid \boldsymbol{\psi}'_{k-1}) p(\boldsymbol{\Psi}'_{k-1} \mid \mathbf{X}_{k-1})}{p(\mathbf{x}_k \mid \mathbf{X}_{k-1})} \tag{3.27}$$

Since the conditional posterior $p(\boldsymbol{\Xi}_k \mid \mathbf{X}_k, \boldsymbol{\Psi}'_k)$ is Gaussian, the first two moments, *i.e.*, the mean and variance, can be computed analytically for a given $\boldsymbol{\Psi}'_k$. Consequently, it becomes possible to use a particle filter to estimate $p(\boldsymbol{\Psi}'_k \mid \mathbf{X}_k)$, while using the Kalman filter to estimate

$p(\boldsymbol{\Xi}_k \mid \mathbf{X}_k, \boldsymbol{\Psi}'_k)$ [70]. Let $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ denote the mean and variance of the Gaussian density $p(\boldsymbol{\Xi}_k \mid \mathbf{X}_k, \boldsymbol{\Psi}'_k)$; each particle is then represented as the set $\{\boldsymbol{\Psi}'^{(i)}_k, \boldsymbol{\mu}^{(i)}_k, \boldsymbol{\Sigma}^{(i)}_k\}$. The Kalman filter equations are used to update the mean and variance, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ [70]:

$$\begin{aligned}
\boldsymbol{\mu}_{k|k-1} &= \mathbf{A}\boldsymbol{\mu}_{k-1} && \text{(3.28a)}\\
\boldsymbol{\Sigma}_{k|k-1} &= \mathbf{A}\boldsymbol{\Sigma}_{k-1}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T && \text{(3.28b)}\\
\mathbf{S}_k &= \mathbf{C}\boldsymbol{\Sigma}_{k|k-1}\mathbf{C}^T + \mathbf{D}\mathbf{D}^T && \text{(3.28c)}\\
\hat{\mathbf{x}}_{k|k-1} &= \mathbf{C}\boldsymbol{\mu}_{k|k-1} && \text{(3.28d)}\\
\boldsymbol{\mu}_k &= \boldsymbol{\mu}_{k|k-1} + \boldsymbol{\Sigma}_{k|k-1}\mathbf{C}^T\mathbf{S}_k^{-1}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) && \text{(3.28e)}\\
\boldsymbol{\Sigma}_k &= \boldsymbol{\Sigma}_{k|k-1} - \boldsymbol{\Sigma}_{k|k-1}\mathbf{C}^T\mathbf{S}_k^{-1}\mathbf{C}\boldsymbol{\Sigma}_{k|k-1} && \text{(3.28f)}
\end{aligned}$$

When it is not possible to sample directly from the marginal posterior distribution $p(\boldsymbol{\Psi}'_k \mid \mathbf{X}_k)$ at time k, an alternative importance distribution is introduced $\pi(\boldsymbol{\Psi}'_k)$ such that [70]:

$$\pi(\boldsymbol{\Psi}'_k) = \pi(\boldsymbol{\Psi}'_{k-1})q(\boldsymbol{\psi}'_k \mid \boldsymbol{\psi}'_{k-1}) \tag{3.29}$$

Using importance sampling, the weights are computed as:

$$\hat{w}_k = \frac{p(\boldsymbol{\Psi}'_k \mid \mathbf{X}_k)}{\pi(\boldsymbol{\Psi}'_k)} \tag{3.30}$$

and using equation (3.27), a sequential weight update is obtained:

$$\hat{w}_k \propto w_{k-1}p(\mathbf{x}_k \mid \boldsymbol{\Psi}'_k, \mathbf{X}_{k-1})\frac{p(\boldsymbol{\psi}'_k \mid \boldsymbol{\psi}'_{k-1})}{q(\boldsymbol{\psi}'_k \mid \boldsymbol{\psi}'_{k-1})} \tag{3.31}$$

where the likelihood function $p(\mathbf{x}_k \mid \boldsymbol{\Psi}'_k, \mathbf{X}_{k-1})$ is given by [70]:

$$p(\mathbf{x}_k \mid \boldsymbol{\Psi}'_k, \mathbf{X}_{k-1}) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{S}_k) \tag{3.32}$$

and $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{S}_k$ are evaluated using the Kalman filter equations listed in equations (3.28). Additionally, due to the dependence on previous values of the state $\boldsymbol{\Xi}_{k-1}$, the likelihood function does not simplify further [70]. An algorithm from the literature which relies on the RBPF for sequential frequency estimation [70] is reproduced in Section 3.5.

3.3 Model Order Selection via jump Markov system and reversible jump MCMC

In certain estimation problems, the number of parameters to be estimated may be one of the parameters itself. For example, in the context of frequency estimation, when not known *a priori*, the number of tones present in the signal is itself a parameter be determined. The number of parameters may be regarded as the *model order*, and model order selection may need to be performed in such cases [30, 32, 86, 87].

A more general class of problem involves estimation of the *regime*, where the regime may correspond to different signal models. For example, one regime may correspond to an aircraft flying at constant velocity (constant velocity model) while another may be used to model the aircraft executing a turn (coordinated turn model) [72] and additionally, each regime may require estimation of a different set of parameters.

In the case of model order estimation, each regime corresponds to a state with a different number of tones. In particular, two methods for model order selection are considered here. The first involves a jump Markov system (JMS) which requires establishment of a set of transition probabilities which govern whether the model order will increase, decrease, or remain unchanged at each iteration. An alternative approach is through the use of a reversible jump MCMC (RJMCMC) method which is a framework which allows comparison of distributions which occur exist over different dimensional spaces. Transition probabilities are used to first propose any change to the model order which is then accepted or rejected according to a certain acceptance criterion. This section provides an overview of both methods, while their particular implementation for estimation of a varying number of tones and harmonics is considered later in Sections 3.5 and 5.2.

3.3.1 Jump Markov Systems

A JMS offers one method for dealing with problems involving changing regimes [72,79,88,89]. In such problems, in addition to a possibly continuous state, there is also a discrete state variable which denotes the regime. While a JMS can in general be used to deal with systems that have different regimes of operation, in the context of sequential frequency estimation, the different regimes could correspond to the presence of different number of frequency tones in the signal, *i.e.*, different model orders.

The JMS provides the framework by which changes between different regimes can be modelled. Let P_k denote the mode at time k, $P_k \in S$, where S represents a discrete, finite set of the possible values of P_k. A Markov chain can be constructed on P_k by specifying the transition probability of the mode from i to j [88] as $p(P_{k+1} = j \mid P_k = i), \{i, j\} \in S$.

If the number of elements in S is denoted by n_S, then it is possible to construct a square matrix of size n_S containing the transition probabilities between all possible modes. This matrix describes all possible changes to the regime variable P_k. The JMS can be used to model the change in the regime P_k and associated state $\boldsymbol{\psi}_k$, which itself may be a function of P_k, and thus estimate the posterior distribution $p(\boldsymbol{\Psi}_k, \mathbf{P}_k \mid \mathbf{X}_k)$, where $\mathbf{P}_k = P_{1:k}$.

As an alternative to JMS, it is possible to use a reversible jump sampler to determine the mode P_k. Application of the JMS to frequency estimation will be addressed in detail in Section 3.5. The following section will introduce RJMCMC methods and discuss their advantages and disadvantages compared to the JMS framework.

3.3.2 Reversible Jump MCMC Methods

Reversible jump methods [30] find widespread application to problems which require model selection or model order selection. In such cases, the dimension of the unknown state (to be estimated) may be variable. The reversible jump sampler allows implementation of model order determination within a Bayesian framework. For the problem addressed in this work, *viz.*, sequential frequency estimation, the reversible jump sampler can be used to determine the number of frequency tones and harmonics. In dealing with the variable model order, the variable dimension of the parameter vector must be accounted for, and the reversible jump sampler offers a suitable means for comparing parameter vectors which possess different dimensions.

The reversible jump MCMC (RJMCMC) sampler is has been used for joint determination of model order and parameters (see, for example [35], where it is used to estimate an unknown number of sources in coloured noise) and has been used to estimate an unknown number of sinusoids in noise [29, 31]. While these problems are addressed within a batch offline scenario, it is possible to use RJMCMC methods within a SMC framework.

While it is possible to perform joint model order estimation using the JMS framework [48, 70], the use of the reversible jumps sampler offers certain benefits. In a JMS, transition probabilities are defined between the different regimes. In contrast, the reversible jump sampler proposes

changes to the mode with a certain probability, but then accepts the new state only if, on average, the new state is associated with a higher probability in the posterior distribution. When used in a particle filter framework, this accept/reject mechanism ensures that good particles are not lost, thus reducing the variance of the weights and limiting particle degeneracy [90]. This benefit, however, comes at the cost of applying the sampler to each particle in the filter.

In contrast with batch offline MCMC methods, the reversible jump sampler does not require a burn-in period when used in the SMC context [34]. Burn-in refers to the practice of discarding the first iterations of an MCMC run so as to ignore samples which are not drawn from the limiting distribution. In the particle filter, the samples are already distributed according to the posterior distribution. Thus, there is no need for a burn-in period to ensure that the samples achieve the desired distribution. As a result of this, the reversible jump sampler can be run for a single iteration and this single iteration is used to achieve model or model order determination [34].

The general structure of the reversible jump sampler as described by Green [30] will be summarised here. The particulars of its application to model order determination for the frequency estimation problem will be addressed in Section 5.3.

As described in the previous section, let S denote the discrete, finite set of candidate regimes. Let $P_k \in S$ denote the regime at time k. For the regime $P = i, i \in S$, an associated state is defined as $\boldsymbol{\omega}(i) \in \mathbb{R}^{n_P}$ where n_P denotes the dimensions associated with the regime P.

A move is proposed from state $\boldsymbol{\psi} = \{P_k = i, \boldsymbol{\omega}(i)\}$ to $\boldsymbol{\psi}^\star = \{P_k = j, \boldsymbol{\omega}(j)\}$ with transition probability $p(\boldsymbol{\psi}^\star \mid \boldsymbol{\psi})$. When the dimensions of the states $\boldsymbol{\psi}^\star$ and $\boldsymbol{\psi}$ are not the same, it becomes necessary to introduce a dummy variable for dimension matching. Discussion of the dimension matching requirement will be carried out in greater detail in Section 5.3.

The new state $\boldsymbol{\psi}^\star$ is then accepted according to the acceptance ratio r, given by [30, 31]:

$$r = \underbrace{\frac{p(\boldsymbol{\psi}^\star \mid \mathbf{x}_k)}{p(\boldsymbol{\psi}_k \mid \mathbf{x}_k)}}_{\text{posterior ratio}} \underbrace{\frac{d(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_k^\star)}{d(\boldsymbol{\psi}_k^\star \mid \boldsymbol{\psi}_k)}}_{\text{proposal ratio}} \underbrace{\mathbf{J}}_{\text{Jacobian}} \tag{3.33a}$$

where the Jacobian is evaluated as:

$$\mathbf{J} = \left| \frac{\partial(\boldsymbol{\omega}(j))}{\partial(\boldsymbol{\omega}(i))} \right|$$

Given the acceptance ratio, the acceptance probability is then evaluated as:

$$\alpha = \min\{1, r\} \tag{3.33b}$$

Through this framework, the RJMCMC sampler allows estimation of parameters when the parameter space spans multiple dimensions. The construction of specific moves, as well as the distributions in equation (3.33a) will be discussed in Section 5.3 once the framework for sequential frequency estimation has been established.

3.4 Methods for Time-Frequency Analysis

The previous chapter examined both non-parametric as well as parametric methods for time-frequency analysis. This section further examines parametric methods for sequential frequency estimation, focusing on those based on Monte Carlo methods. The first portion examines methods in the existing literature before examining alternative ideas in Section 3.4.2. Following this, the chapter concludes with a detailed description of a sequential frequency estimation algorithm developed by Dubois and Davy in [70].

3.4.1 Established methods for time-frequency analysis

This section discusses various SMC methods for sequential time-varying frequency estimation. The spectrum can be estimated using either a time-varying autoregressive (TVAR) or harmonic model. The discussion begins with a review of a method developed by Andrieu *et al*. [91, 92] using an AR model.

Tracking TVAR parameters: Andrieu *et al*., in [91, 92], develop a method which relies on a particle filter to track the evolution of the poles in an all-pole spectrum over time. The authors adopt a model where the AR coefficients vary across frames, but not within them (where a frame is defined as a short sequence of observations). In [91], the authors adopt a fixed model

order, *i.e.*, a model with a fixed number of AR coefficients. This model is extended in [92] where the authors propose a model in which the number of poles in the AR model is not fixed and a JMS is used to determine the time-varying number of poles. The observed signal y_t, at time t, is modelled as:

$$y_t = \sum_{i=1}^{K_t} a_t^{(i)} y_{t-k} + e_t \tag{3.34}$$

where K_t is the number of time-varying AR coefficients, $a_t^{(i)}$ is the i^{th} AR coefficient, and e_t is zero-mean, white Gaussian noise with variance $\sigma^2_{(e)}$. The number of AR coefficients K_t also defines the length of the frame and the observation samples prior to y_{t-K_t} are ignored.

Rather than tracking the AR coefficients directly, the authors track the poles in the spectrum and define a non-linear transformation from the poles to the AR coefficients. The transformation from the set of instantaneous frequencies ν_t, their moduli ρ_t, real poles $\mathbf{r}_t$ and complex conjugate poles $\mathbf{z}_t$ onto the set of AR coefficients is defined as:

$$\mathbf{a}_t = \Phi(\nu_t, \rho_t, \mathbf{r}_t, \mathbf{z}_t) \tag{3.35}$$

such that the observation equation is now written as:

$$y_t = \Phi(\nu_t, \rho_t, \mathbf{r}_t) y_{t-1:t-K_t} + e_t. \tag{3.36}$$

A JMS is used to model evolution of the time-varying number of poles K_t. A birth move increases the number of poles, a death move reduces the number, and an update-only move leaves the number of poles unaltered. These moves are specified using a set of predefined transition probabilities and there is no accept-reject mechanism for rejecting bad proposal moves. The transition probabilities dictate whether, given the current number of poles, a new pole should be added or an existing one deleted. Further discussion of JMS and RJMCMC methods for determination of a time-varying number of tones is presented in Sections 3.5 and 5.3 respectively. The authors subsequently apply the algorithm and demonstrate estimation of the frequency content of a time-varying signal.

Tracking ridges in the STFT: As mentioned in the previous chapter, tracking ridges in a time-frequency representation can provide an estimate of the constituent frequencies of a signal. Such an approach is followed by Dubois *et al*. in [48] and [49] where the algorithm involves

using a particle filter to sequentially track the peaks in the STFT sequentially. The use of a particle filter allows implementation of a tracking algorithm in the presence of a non-linear observation model which translates a set of frequencies into the frequency observation space by examining the short time spectrum. The signal is windowed in the time-domain to obtain a local estimate of the k^{th} frame of the signal as:

$$x_{k,\tau} = x_{k+\tau-\frac{L}{2}} w_\tau, \quad k \in \{1, \ldots, T\}, \tau \in \{1, \ldots, L\} \tag{3.37}$$

where the window, or frame, length is specified as L, the total length of the signal is T and $w_\tau, \tau \in \{1, \ldots, L\}$ denotes the value of the window function at time τ. The observed sample is further modelled as:

$$x_{k,\tau} = s_{k,\tau} + b_\tau \tag{3.38}$$

where v_τ is additive white Gaussian noise, and $s_{k,\tau}$ is modelled as the sum of P_k sinusoids where the p^{th} sinusoid has amplitude $a_k^{(p)}$ and frequency $f_k^{(p)}$:

$$s_{k,\tau} = w_\tau \sum_{p=1}^{P_k} a_k^{(p)} \cos(2\pi f_k^{(p)} \tau). \tag{3.39}$$

Let $s_{k,\tau} \rightleftharpoons S_{k,\tau}$ denote the Fourier transform pair of the windowed observation sequence. Then, the observation space is constructed as a non-linear operation on the signal $x_{k,\tau}$ to yield:

$$y_{k,\tau} = |S_{k,\tau}| + e_{k,\tau} \tag{3.40}$$

where $|S_{k,\tau}|$ denotes the magnitude of the Fourier transform and $e_{k,\tau}$ is zero-mean, white Gaussian noise with variance $\sigma^2_{(e),k}$. The use of the non-linear transformation, $s_{k,\tau} \rightarrow y_{k,\tau} \approx S_{k,\tau}$, allows for estimation of the magnitude and frequency of the tones without needing to consider the phase of the signal which can be ignored. A drawback of this approach is its reliance on the STFT, which can, in certain cases, provide an ambiguous time-frequency representation of the signal and the method is thus constrained by the limitations of the STFT. This will be illustrated in the following section when bat calls are considered for analysis.

The aim of the algorithm is, using the STFT, to estimate the time-varying parameters P_k, $\mathbf{f}_k = [f_k^{(1)}, \ldots, f_k^{(P_k)}]^T$ and $\mathbf{a}_k = [a_k^{(1)}, \ldots, a_k^{(P_k)}]^T$ which denote the number of tones, their frequencies, and amplitudes respectively. As in the case of the previously discussed algorithm by Andrieu *et al*. [91, 92], a JMS is used to estimate the number of tones using birth, death and

update-only moves. The birth and death moves in this case, correspond to addition and removal of tones respectively, while an update-only move leaves the number of tones unchanged. Additionally, to allow for dynamics in the variation of the parameters, a random walk is specified for the frequency terms:

$$f_{k+1}^{(p)} = f_k^{(p)} + v_{(f),k}^{(p)} \tag{3.41}$$

such that the process noise $v_{(f),k}^{(p)} \sim \mathcal{N}(0, (\sigma_{(f),k}^{(p)})^2)$ possesses a time-varying variance which evolves as:

$$\left(\sigma_{(f),k+1}^{(p)}\right)^2 = \left(\sigma_{(f),k}^{(p)}\right)^2 e^{\phi_k^{(p)}}, \quad \phi_k^{(p)} \sim \mathcal{N}(0, \sigma_\phi^2) \tag{3.42}$$

or equivalently, the logarithm of the variance $\left(\sigma_{(f),k+1}^{(p)}\right)^2$ evolves according to a random walk ensuring that the variance is always positive. A similar parameterisation is set up for the time-varying amplitude parameters as well as the evolution of the observation noise variance $\sigma_{(e),k}^2$.

This system of equations can be fit into a particle filtering framework to sequentially estimate the frequencies and amplitudes of the signal. To improve the performance of the filter, the authors propose a variation of the basic particle filter which uses the unscented transform [74] to construct the proposal distribution.

The unscented transform, which gives rise to the unscented Kalman filter (UKF), propagates a set of deterministic samples (known as sigma points) through the non-linear state space equations such that the mean and covariance of the posterior distribution can be recovered correctly. When used in a particle filter framework, this resulting filter is referred to as the unscented particle filter [93].

The unscented transform improves performance by constructing a sampling distribution which is closer to the optimal sampling distribution and increases the rate of convergence in cases where the prior probability is centred in the tails of the likelihood function. Using the unscented transform, the authors show that acceptable performance can be achieved even when a large number of particles is not used. Dubois and Davy [70] have also developed an alternative sequential frequency estimation algorithm which does not rely on tracking spectral peaks in the STFT. This method is discussed in Section 3.5.

Tracking of piecewise-linear chirps: A slightly different approach is adopted for IF estimation in Li *et al*. [94]. In that method, a time-varying frequency is approximated as a piecewise-linear chirp, and the aim is to estimate the number of components, as well as their parameters.

The authors utilise a particle filtering algorithm for estimation of static parameters [95].

Lee and Chia [95] note that particle filters developed for tracking dynamic states are unable to cope with unknown static parameters. Their presence "violates the mixing properties required for [particle filter] algorithms" [95] resulting in instability in the particle filter algorithm. A solution proposed by both Lee and Chia [95], as well as Chopin [96], is to perform rejuvenation of the particles through an additional MCMC step to improve the mixing properties. The Monte Carlo rejuvenation step consists of applying an independent Metropolis-Hastings sampler [36] which proposes values around the current estimate of the parameters while the remaining steps in the particle filtering algorithm remain unchanged.

The algorithm developed by Li *et al*. [94] attempts to fit all the time samples up to the current time using a (set of) linear chirp(s). The change in the posterior distribution from the initial/start time up to the current time instant is measured using the Kullback-Leibler [97] distance, and when this exceeds a threshold, it indicates that the current observations can no longer be represented using the current static parameters. In this case, a Monte Carlo rejuvenation step is carried out and the algorithm attempts to fit the observations from the current time instant onwards using a new set of chirp parameters. The authors demonstrate the method using FM signals and subsequently incorporate aspects of chirp model selection.

This section examined methods from the literature for estimation of a time-varying spectrum. The following section examines different approaches to the analysis of echolocation calls. The purpose is to show the limitations of certain approaches, and in doing so, illustrate the reasoning behind the course taken in this work.

3.4.2 Time-frequency analysis of bat echolocation signals

To understand the motivation behind the approach adopted in this work, this section considers the analysis of bat echolocation calls. Three calls are considered here, from *M. nattereri*, *P. pipistrellus* and from *Myotis daubentonii* (Kuhl, 1817). Figures 3.4-3.6 show each of these calls along with their spectrograms, where the sampling rate has been normalised to 1 Hz. The spectrogram is capable of showing the relevant detail present in the calls in most cases. However, on inspection of the call from *M. daubentonii*, ambiguities in the STFT are clearly visible (this signal will be studied further in Chapter 7). This suggests that merely the STFT or non-parametric methods are not sufficient to analyse the signals. Additionally, as discussed previously, using

non-parametric methods for analysis would entail some form of post-processing to extract the call from the time-frequency representation.

The use of a parametric approach helps in circumventing this problem by either estimating the precise frequencies that are present in the signal, or by fitting the signal to a particular chirp and estimating its parameters. For example, this is illustrated in Figure 3.7 where the parameters that describe the particular chirp need to be determined. In this case, the start and end times and frequencies need to be determined as well as the parameters defining the frequency modulation law of the chirp. Chirp parameter estimation will be considered here for a call from Natterer's bat and a pipistrelle using a linear and exponential chirp respectively.

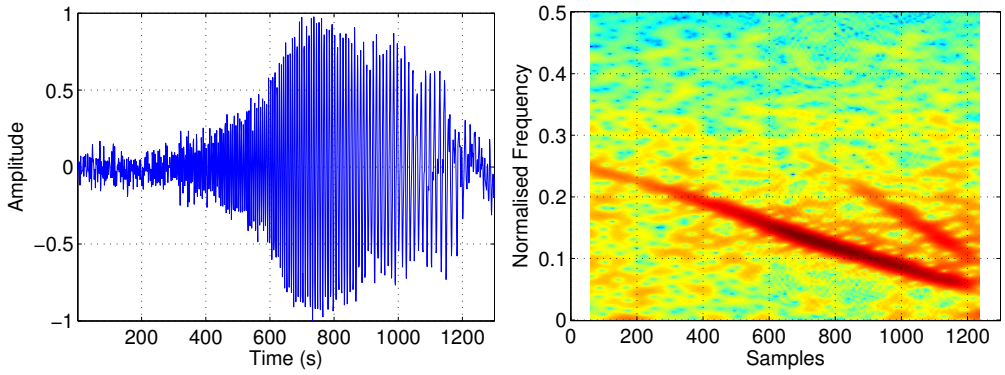
The call from *M. nattereri* is considered first. Neglecting higher harmonics, the call is assumed to fit a linear chirp of the form:

$$y_t = a\cos\left[2\pi\left(f_{start} + \frac{c}{2}t\right)t\right] + b\sin\left[2\pi\left(f_{start} + \frac{c}{2}t\right)t\right], \quad t \in \{0, \ldots, T\} \tag{3.43}$$

where f_{start} and f_{end} respectively denote the initial and final frequency of the chirp, T represents the duration of the chirp, and the chirp rate is given by $c = \dfrac{f_{end} - f_{start}}{T}$. Since the model is linear in the amplitudes (a, b), and assuming additive white Gaussian noise, it is possible to marginalise the amplitudes and noise variance as illustrated in [17, 62] and in Section 4.4.4. Using this, the posterior distribution $p(f_{start}, f_{end} \mid \mathbf{Y}_T)$ is evaluated on a grid parameterised on the start and end frequencies of the linear chirp.

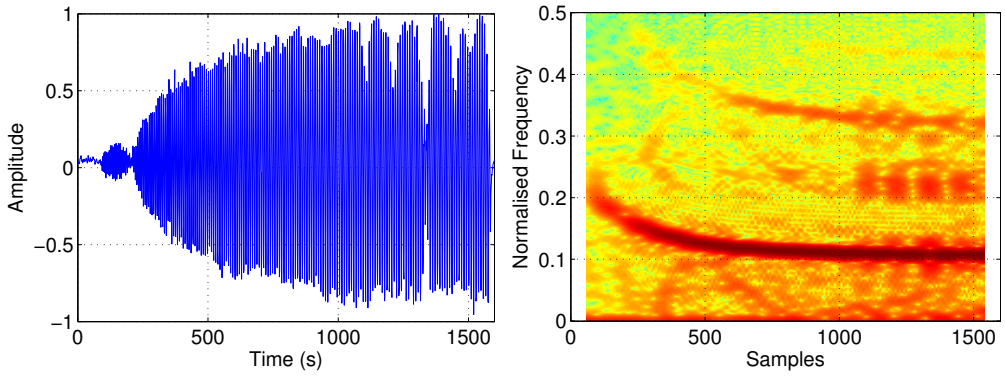
The call used in this example has been manually selected from a recording such that the segment corresponds to a single call. Figure 3.8 shows the log posterior distribution over the normalised start and end frequencies for the call. By evaluating the posterior, it is possible to obtain parameter estimates for the call considered by searching for the global maximum. Selecting the maximum from the distribution yields a chirp start frequency of 0.2660 and a stop frequency of 0.0375. These estimates are limited by the grid resolution, which is 0.0005 in this example.

An examination of the posterior distribution shows an extent of blurring, and the maximum is seen to lie somewhere along a ridge. This occurs since the linear chirp is not perfectly matched to the call, and an examination of the spectrogram of the call reveals a minor deviation from a linear chirp structure. Such deviations can make it difficult to accurately determine the parameters of a call.



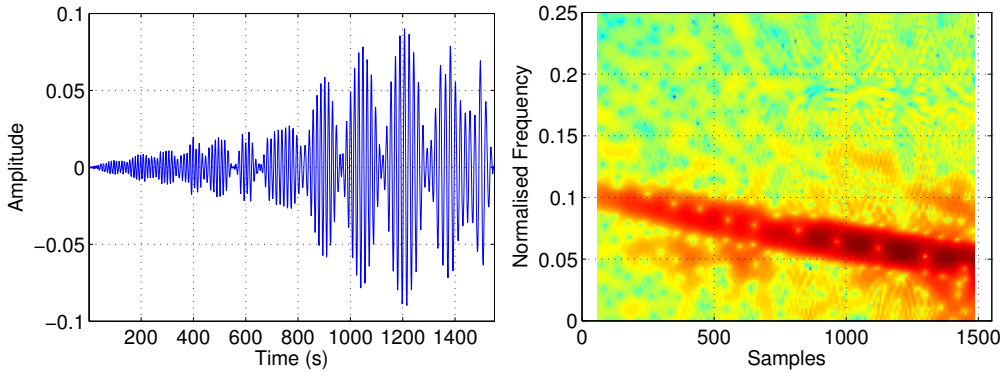
(a) (b)

Figure 3.4: *a) Example of a call from* M. nattereri *and b) the spectrogram of the call.*



(a) (b)

Figure 3.5: *a) Example of a call from* P. pipistrellus *and b) the spectrogram of the call.*



(a) (b)

Figure 3.6: *a) Example of a call from* M. daubentonii *and b) the spectrogram of the call.*

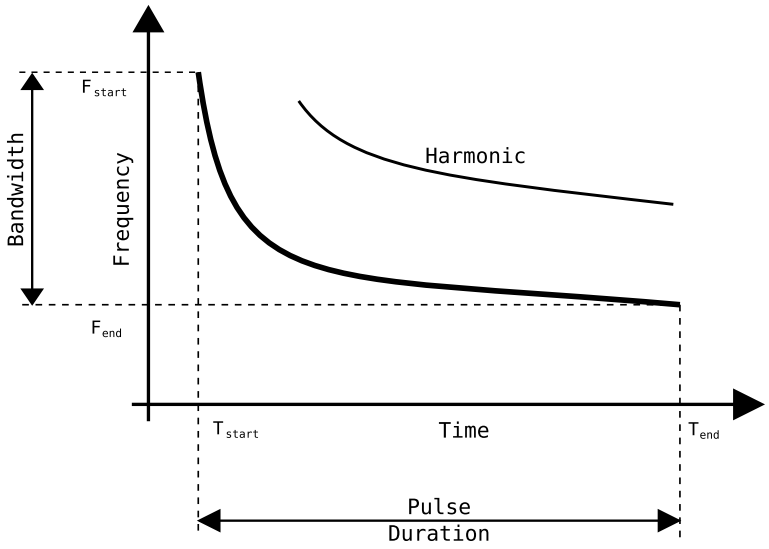


Figure 3.7: *Example of chirp parameters which may need to be estimated.*

Monte Carlo methods can be used to estimate the chirp parameters rather than evaluating the posterior distribution over a grid. The following example adopts an approach similar to that of Li *et al*. [94] and uses a particle filter to estimate static chirp parameters. In specifying the parameters to be static, the particle filter is required to sequentially refine the parameter estimates describing the chirp structure as more data becomes available. As discussed earlier, since static parameters are present in the signal model, following the method of Lee and Chia [95] and Chopin [96], a Monte Carlo rejuvenation step needs to be included in the particle filter to ensure stability.

This algorithm is applied to estimate the chirp parameters of the two calls. Figures 3.9 and 3.10 show the sequential estimates of the initial frequency and the chirp rate obtained from the filter. From the figure, the chirp start frequency is estimated to be 0.2567 Hz, while the stop frequency is 0.0339 Hz. While the figures show that the algorithm eventually converges to frequency and chirp rate values after a certain number of iterations, it is seen that there is a mismatch between these estimates and those previously obtained from a grid search of the posterior.

A similar result is obtained for the call from *P. pipistrellus* shown in Figure 3.5. As discussed in Section 2.4, several chirp models have been used to fit echolocation calls. In the example considered here, the call is fit to an exponential chirp given by equation (2.54), reproduced here:

$$f_t = B \exp\left(\frac{t}{T_D}\right) + f_a \tag{3.44}$$

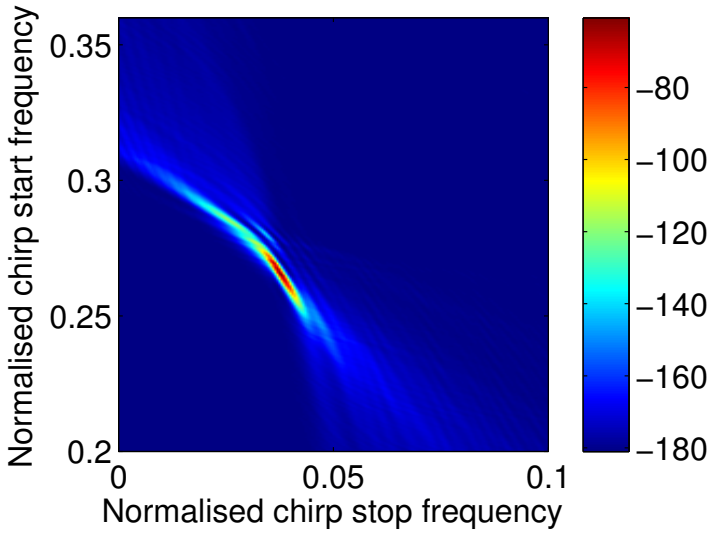


Figure 3.8: *Log posterior distribution over the initial and final frequencies from fitting the call from* M. nattereri *to a linear chirp.*

where T_D is the time decay constant (negative), f_a is the asymptotic frequency and B is the bandwidth of the chirp. Fitting the call to the exponential chirp results in obtaining parameter estimates as shown in Figures 3.11-3.13.

Despite the ability of the algorithm to arrive at parameter estimates, several drawbacks, are associated with the chirp parameter estimation approach to echolocation call analysis. Firstly, while the calls have been manually selected in this example, a general case would involve parsing and selecting individual calls from a longer duration recording, an aspect which would need to be automated. While this is not insurmountable, it may result in clipping of the start or end of a call, depending on the method adopted.

More important, however, is the limitation of the models used in the parameter estimation exercise. The models used here assume a constant amplitude envelope of the call when it is clear that the calls are amplitude modulated. Additionally, the models do not account for harmonics in the call. Estimation of the harmonics is further complicated by the fact that they occur only over part duration of the call. The models described in the previous chapter may not adequately represent the significant call variety which is observed in what is a very large number of species of bats. For example, a linear chirp may offer only a poor approximation of the call from Daubenton's bat shown in Figure 3.6 which results in poor characterisation of a

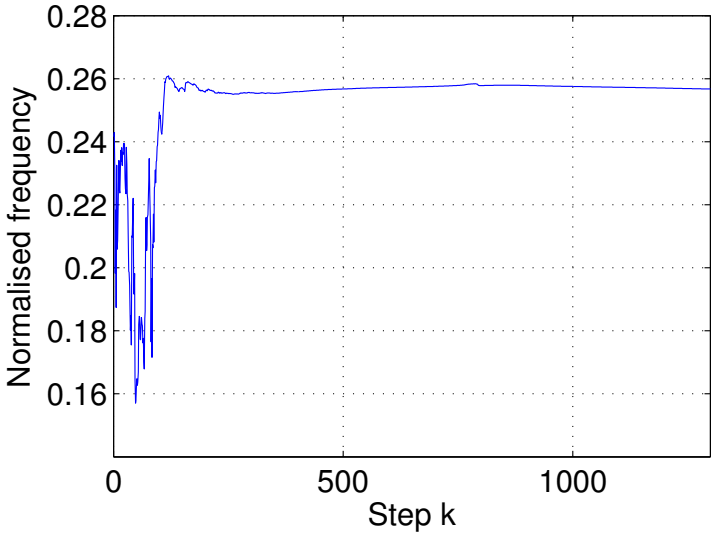


Figure 3.9: *Sequential estimate of the initial frequency of the call from* M. nattereri *when fit to a linear chirp model.*

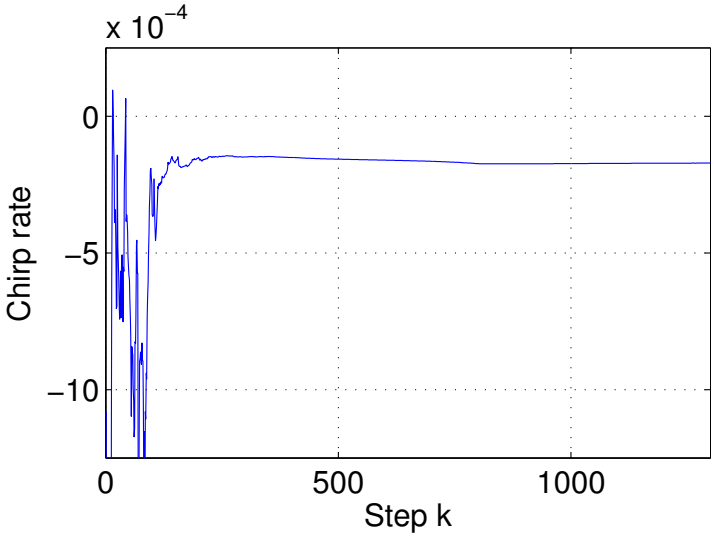


Figure 3.10: *Sequential estimate of the chirp rate of the call from* M. nattereri *when fit to a linear chirp model.*

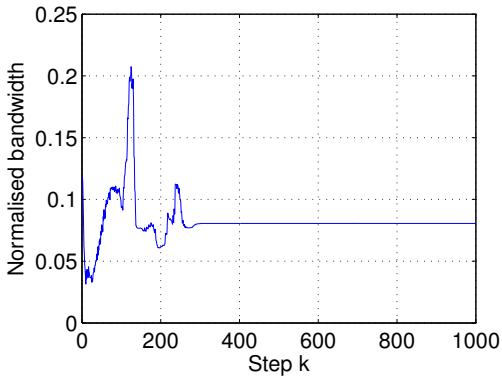


Figure 3.11: *Sequential estimate of the bandwidth B of the call from* P. pipistrellus.

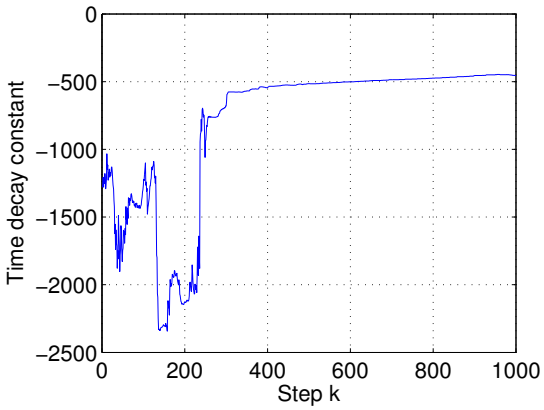


Figure 3.12: *Sequential estimate of the time decay constant T of the call from* P. pipistrellus.

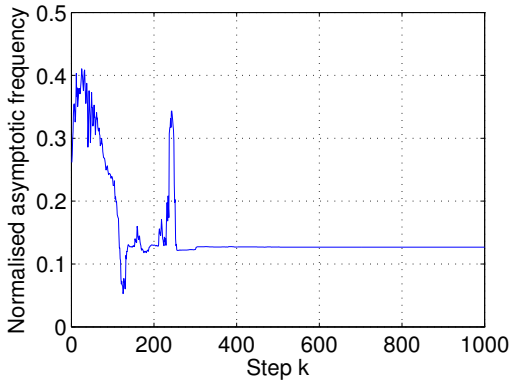


Figure 3.13: *Sequential estimate of the asymptotic frequency f_a of the call from* P. pipistrellus.

peak in the posterior distribution.

Consequently, a more prudent approach is to first estimate the time-varying frequencies present in the call. By explicitly accounting for a varying number of tones and harmonics, as well as the frequency and amplitude modulated nature of the tones themselves, it may be possible to arrive at a more accurate representation of the calls. Should the need arise, chirp parameter estimation can then be performed on this time-frequency representation, however, it is only the frequency estimation aspects that will be considered in the following chapters.

The following section introduces a method from the existing literature for sequential detection and estimation of a time-varying number of tones. This method will be compared with a novel method developed later in Chapter 5.

3.5 Frequency Estimation using the RBPF

Dubois and Davy [70] develop a framework for joint detection and estimation of time-varying sinusoids within a particle filter framework. A JMS is used to describe the evolution of the model and a RBPF is used to achieve marginalisation of the amplitude parameters. The implementation of this method (used in the comparison in Section 5.5) contains some deviations from the original method described in [70]. In summarising the algorithm, these deviations will be made explicit.

3.5.1 Signal Model

The signal model used in [70] is a particular case of the harmonic model described in Chapter 4. While the full model is developed in that chapter, the specific instance of the model used in this algorithm will be described here.

The model used by Dubois and Davy uses a sinusoidal basis to represent a short segment of the signal. The signal is composed of a set of fundamental tones, and each tone is associated with a fixed, large number of harmonics. The frequency content is estimated within a sliding window of length L, and the subscript k is used to denote the frame index. The windowed signal (within the frame) is centred about time $t_k \in \tau = \{0, \ldots, T-1\}$, where T is the total length of the signal. The frequency content of the k^{th} frame is evaluated over the segment $l_k \in \{t_k - L', \ldots, t_k + L'\}$, where $L' = (L-1)/2$. The model within the k^{th} frame is then

written as [70]:

$$x_{l_k} = \sum_{p=1}^{P_k} \sum_{h=1}^{H} \left(a_k^{(p,h)} \cos \phi_{k,l_k-t_k}^{(p,h)} + b_k^{(p,h)} \sin \phi_{k,l_k-t_k}^{(p,h)} \right) + e_{l_k} \tag{3.45a}$$

where

$$\phi_{k,l}^{(p,h)} = 2\pi h f_k^{(p)} l \tag{3.45b}$$

In the model, H is the (constant and large) number of overtones associated with each tone. Using a lowpass assumption, the authors note that the true number of overtones may be less than H since the frequency of the overtones may not exceed Nyquist frequency. By specifying H sufficiently large, the system is over-modelled. The amplitudes of the cosine and sine contributions of the h^{th} harmonic of the p^{th} tone are respectively given by $a_k^{(p,h)}$ and $b_k^{(p,h)}$. As a result of the over-modelling, amplitude estimates of non-existent harmonics will be negligible and these can be removed from the final estimate by setting a threshold for the amplitude terms. The observation noise is zero-mean Gaussian of fixed variance $e_{l_k} \sim \mathcal{N}(0, \sigma_{(e)}^2)$.

3.5.2 Likelihood Function and Prior Distributions

The vector of parameters to be determined is denoted by $\boldsymbol{\psi}_k = \{P_k, \mathbf{f}_k, \mathbf{a}_k\}$, where $\mathbf{f}_k = [f_k^{(1)}, \ldots, f_k^{(P_k)}]^T$ and $\mathbf{a}_k = [a_k^{(1,1)}, \ldots, a_k^{(P_k,H)}, b_k^{(1,1)}, \ldots, b_k^{(P_k,H)}]^T$ are vectors of frequencies and amplitudes respectively. Additionally, the observation and noise vectors are given by $\mathbf{x}_k = [x_{t_k-L'}, \ldots, x_{t_k+L'}]^T$ and $\mathbf{e}_k = [e_{t_k-L'}, \ldots, e_{t_k+L'}]^T$ respectively. To arrive at an expression for the likelihood function, the following matrices are defined:

$$\boldsymbol{\Phi}_k = \begin{bmatrix} \phi_{k,-L'}^{(1,1)} & \cdots & \phi_{k,-L'}^{(1,H)} & \phi_{k,-L'}^{(2,1)} & \cdots & \phi_{k,-L'}^{(P_k,H)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \phi_{k,L'}^{(1,1)} & \cdots & \phi_{k,L'}^{(1,H)} & \phi_{k,L'}^{(2,1)} & \cdots & \phi_{k,L'}^{(P_k,H)} \end{bmatrix} \tag{3.46a}$$

$$\mathbf{G}_k = [\cos(\boldsymbol{\Phi}_k) \quad \sin(\boldsymbol{\Phi}_k)] \tag{3.46b}$$

where $\cos(\cdot)$ and $\sin(\cdot)$ operate element-wise on $\boldsymbol{\Phi}_k$. The signal model can then be rewritten in terms of these matrices as:

$$\mathbf{x}_k = \mathbf{G}_k \mathbf{a}_k + \mathbf{e}_k \tag{3.47}$$

where $\mathbf{e}_k = [e_{t_k-L'}, \ldots, e_{t_k+L'}]$. For a signal model which can be written in this form, the likelihood function can be written as [17, 62]:

$$p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) = \frac{1}{(2\pi\sigma^2_{(e)})^{\frac{L}{2}}} \exp\left[-\frac{\|\mathbf{x}_k - \mathbf{G}_k\mathbf{a}_k\|^2}{2\sigma^2_{(e)}}\right] \tag{3.48}$$

The construction and nature of the $\mathbf{G}_k$ matrix will be discussed in greater detail in Section 4.4 where the *general linear model* is introduced.

For the prior distributions, Dubois and Davy [70] specify a flexible framework for the update of the frequency and the amplitude parameters, by using smoothing functions to propose the updates, although only a simple random walk is used in the sequential frequency estimation algorithm here. Additionally, the process noise variance is further parameterised such that the process noise variance is a time-varying parameter controlled by a set of hyper-parameters. This additional parameterisation allows the method to deal with greater variations in the signal dynamics. However, the implementation in this thesis uses a fixed process noise for the parameter update.

The update for the frequency tones is specified as a random walk:

$$f_k^{(p)} = f_{k-1}^{(p)} + v_{(f),k} \tag{3.49}$$

with $v_{(f),k} \sim \mathcal{N}(0, \sigma^2_{(f)})$. Although a simple random walk is not the best choice of model, it is used here due to its simplicity. A more appropriate model for the frequency evolution should be used in practice by incorporating prior knowledge of signal structure into the model. The update to the number of frequency components is specified as [70]:

$$P_k = P_{k-1} + \begin{cases} 1, & \text{with probability } b_k \\ -1, & \text{with probability } d_k \\ 0, & \text{with probability } u_k = 1 - (b_k + d_k) \end{cases} \tag{3.50}$$

where b_k, d_k and u_k represent respectively birth, death and update-only of tones. The probabilities for the update to the number of tones are reproduced from [70] in Table 3.1. It is noted that the probabilities associated with the birth and death processes are small relative to the update-only step. The effect of a large probability associated with the update-only step serves to prevent a large number of particles from moving to spaces which may be associated with low

	$P_{k-1} = P_{min}$	$P_{min} < P_{k-1} < P_{max}$	$P_{k-1} = P_{max}$
b_k	0.1	0.1	0.0
u_k	0.9	0.8	0.9
d_k	0.0	0.1	0.1

Table 3.1: *JMS transition probabilities for birth, death and update moves.*

likelihoods.

The birth and death processes involve the addition and removal of a single tone to/from the vector $\mathbf{f}_k$. Dubois and Davy utilise an efficient sampling distribution which proposes new tones by estimating the dominant components from the observed signal. In the implementation used in this thesis, a simple uniform prior is used for the sampling distribution. One of the aims of Section 5.5 is to compare the performance of the different particle filter marginalisation schemes, and restricting both algorithms to identical (and simple) sampling schemes should have no impact on the performance of the algorithms themselves. In the case of a death move, a tone is selected at random from the vector $\mathbf{f}_k$ and removed.

3.5.3 JMS Framework and State Update Equations

The update of the state parameters $\boldsymbol{\psi}_k$ is given by:

$$
\begin{aligned}
P_k &\sim p(P_k \mid P_{k-1}) && \text{(3.51a)}\\
\mathbf{f}_k &\sim p(\mathbf{f}_k \mid \mathbf{f}_{k-1}) && \text{(3.51b)}\\
\mathbf{a}_k &= \mathbf{A}_k \mathbf{a}_{k-1} + \mathbf{B}_k \mathbf{v}_{(\mathbf{a}),k} && \text{(3.51c)}\\
\hat{\mathbf{x}}_k &= \mathbf{C}_k \mathbf{a}_k + \mathbf{D}_k \mathbf{e}_k && \text{(3.51d)}
\end{aligned}
$$

where the expressions for $\mathbf{a}_k$ and $\hat{\mathbf{x}}_k$ are of the same form as equations (3.24b) and (3.24c) on page 48. The matrices $\mathbf{A}_k$, $\mathbf{B}_k$, $\mathbf{C}_k$ and $\mathbf{D}_k$ are defined as follows [70]:

$$
\begin{aligned}
\mathbf{A}_k &= \mathbb{I}_{2P_kH}\\
\mathbf{B}_k &= \sigma_{(a)}\mathbb{I}_{2P_kH}\\
\mathbf{C}_k &= \mathbf{G}_k\\
\mathbf{D}_k &= \sigma_{(e)}\mathbb{I}_L
\end{aligned}
\qquad (3.52)
$$

where $\mathbb{I}_Q$ denotes the $Q \times Q$ identity matrix. The update for the amplitude parameters is specified as the Gaussian distribution with variance $\sigma^2_{(a)}$, $p(a_k|a_{k-1}) = \mathcal{N}(0, \sigma^2_{(a)})$. The matrix $\mathbf{C}_k$ is identical to the matrix $\mathbf{G}_k$ defined in equations (4.24a) and (4.24b). The $\mathbf{G}_k$ matrix (discussed in the following chapter in the context of the general linear model) is a function of $\{\phi^{(p,h)}_{k,l}\}$, given by equation (3.45b), such that each column of the matrix constitutes a single basis function. Using these definitions and distributions, the RBPF can now be described.

3.5.4 Particle Filtering Algorithm

Dubois and Davy take advantage of the conditional Gaussian state space to implement a solution to the frequency tracking problem using the Rao-Blackwellised particle filter (RBPF). The parameters are partitioned as $\boldsymbol{\psi}'_k = \{P_k, \mathbf{f}_k\}$ and $\boldsymbol{\xi}_k = \mathbf{a}_k$, and the noise variance $\sigma^2_{(e)}$ is assumed known.

The sampling distribution on the $\boldsymbol{\psi}'_k$ parameters is implemented as a random walk as described using equations (3.49) and (3.50). Using the definitions of $\mathbf{A}_k$, $\mathbf{B}_k$, $\mathbf{C}_k$ and $\mathbf{D}_k$ in equation (3.52), the Kalman filter equations, given by (3.28a)–(3.28f), can be evaluated, which allows evaluation of the likelihood function $p(\mathbf{x}_k|\boldsymbol{\Psi}'_k, \mathbf{X}_{k-1})$ using equation (3.32). This allows evaluation of the importance weights using equation (3.31) giving rise to the set of weighted particles as $\{({\boldsymbol{\Psi}'_k}^{(i)}, \mu^{(i)}_k, \boldsymbol{\Sigma}^{(i)}_k), w^{(i)}_k\}, i \in \{1, \ldots, N\}$. Algorithm 3, adapted from [70], lists the steps involved in the RBPF.

3.6 Summary

This chapter introduced SMC methods and their application to the particular problem of time-frequency analysis. The basis of particle filters is provided by SIS, an extension of the importance sampling principle to the sequential case. SIS is used to approximate the posterior distribution using a discrete set of weighted samples. Parameter estimates can be obtained by taking expectations of this posterior distribution which approaches the true estimate as the number of samples, or particles, approaches infinity. In practice, particle filters can provide good results even with a limited number of particles.

Particle filters are seen to suffer from degeneracy when all but a few particles have zero weight. This occurs due to the use of a limited number of particles and can be prevented by using

Algorithm 3 Sequential frequency estimation using the RBPF

Initialisation ($k = 0$):

$(P_0^{(i)}, \mathbf{f}_0^{(i)}) \sim p(P, \mathbf{f}) \mid_{i=1}^{N}$

$\mu_k^{(i)} = \mu_0^{(i)} \mid_{i=1}^{N}$

$\mathbf{\Sigma}_k^{(i)} = \mathbf{\Sigma}_0^{(i)} \mid_{i=1}^{N}$

$w_k^{(i)} = \frac{1}{N} \mid_{i=1}^{N}$

where i is the particle index, and N is the number of particles used.

1: **for** $k = 1 : k_{max}$ **do**

2: Process-update:

$$P_k^{(i)} \sim p(P_k \mid P_{k-1}^{(i)}) \mid_{i=1}^{N}$$
$$\mathbf{f}_k^{(i)} \sim p(\mathbf{f} \mid \mathbf{f}_{k-1}^{(i)}) \mid_{i=1}^{N}$$

Update $\mu_k^{(i)}$ and $\mathbf{\Sigma}_k^{(i)}$ using equations (3.28a)–(3.28f).

3: Evaluate importance weight using equations (3.31) and (3.32):

$$\hat{w}_k^{(i)} \propto w_{k-1}^{(i)} p(\mathbf{x}_k \mid \mathbf{\Psi}_k'^{(i)}, \mathbf{X}_{k-1}) \frac{p(\boldsymbol{\psi}_k'^{(i)} \mid \boldsymbol{\psi}_{k-1}'^{(i)})}{q_k(\boldsymbol{\psi}_k'^{(i)} \mid \boldsymbol{\psi}_{k-1}'^{(i)})}$$

4: Normalise weights:

$w_k^{(i)} = T^{-1} \hat{w}_k^{(i)} \mid_{i=1}^{N}$ where $T = \sum_{i=1}^{N} \hat{w}_k^{(i)}$

5: Obtain parameter estimates (equations (5.17)-(5.19)).

6: Resample:

$[\boldsymbol{\psi}_k'^{(i)}, w_k^{(i)}] = \text{RESAMPLE}[\{\boldsymbol{\psi}_k'^{(i)}, w_k^{(i)}\}_{i=1}^{N}]$

7: **end for**

a resampling step which replaces those particles having low weights with those having high weights. This can be understood as moving particles from low likelihood regions to higher likelihood regions.

Section 3.2 discusses marginalisation of parameters within the particle filter framework. The marginalisation process serves to reduce the dimensions of the parameter space and also serves as a variance reduction method. The resulting filter, called the RBPF, separates the state space such that one set of parameters can be represented using a conditional Gaussian distribution. The observation equation is linear in the amplitudes and thus, the distribution on the amplitudes can be written as a Gaussian distribution conditional on the remaining parameters. The RBPF then relies on the Kalman filter to estimate this distribution while using the particle filter to estimate the remaining non-linear parameters, achieving marginalisation of the amplitude parameters in this case.

Section 3.4 discusses several algorithms for time-frequency analysis which rely on a Monte Carlo framework. The section also discussed methods for chirp parameter estimation of bat echolocation calls while noting that such an approach is hindered by several shortcomings

which arise from a significant amount of variability in bat echolocation calls.

A sensible approach to the analysis of bat echolocation calls is thus, to estimate the time-varying number number of components sequentially. In doing so, the constraint of accounting for the entire structure of the call is lifted, requiring only that the model take into account non-stationarity on a local scale, and under the right circumstances, the signal may be treated as locally stationary.

The concluding section of this chapter introduced a method developed by Dubois and Davy [70] which estimates the time-varying tones of a signal sequentially while assuming the signal is locally stationary. The method developed by Dubois and Davy relies on a JMS to determine the time-varying number of tones through predefined transition probabilities and uses the RBPF to achieve parameter marginalisation.

Before developing the proposed particle filtering algorithm, the following chapter introduces the signal model used for the task. Due to the highly variable nature of bat echolocation calls, the signal model needs to be sufficiently flexible, capable of dealing with frequency modulation as well as a variable number of harmonics. Following the introduction of the model, the relevance of the model to bat echolocation calls will also be examined in the following chapters.

Chapter 4
A Signal Model for Sequential Frequency Estimation

Introduction

Parametric approaches to signal analysis rely on exploiting signal structure to produce a more accurate estimate of the signal parameters. While dealing with the problem of sequential frequency estimation, it is the analysis of bat echolocation call structure that is of particular relevance in this work. Much of the analysis of calls relies on using the STFT to obtain a time-frequency representation of the signals followed by some form of post-processing; for example, Obrist *et al*. use the spectrogram of calls to perform species identification and classification [98].

While the STFT is a powerful analysis tool, as illustrated in the previous chapter, a drawback of the method is its inability to characterise detailed signal structure, occasionally leading to ambiguous time-frequency representations. This motivates the need for an alternative method which facilitates easy analysis of echolocation calls.

The following sections introduce the model used in this work for analysis of echolocation calls while discussing the motivation for the choice of model. The model described in this chapter combines several models from existing literature to arrive at a model which may be suitable for the analysis of echolocation calls.

4.1 A Parametric Model for Bat Echolocation Signals

A parametric approach involves fitting a model to the observed signal. The problem then reduces to determining the parameters of that model which would produce the observed data. For example, as discussed in the previous chapter, time-varying spectral estimation in several methods consisted of estimation of a set of AR coefficients or frequencies (which constitute the basis for a harmonic model).

The choice of model stems from the structure of the signals of interest. Bat echolocation calls are frequency and amplitude modulated signals which do not contain any impulse-type components. Depending on the species, bats produce different types of chirp signals (see Section 2.4 for an overview). In order to analyse individual calls, they must first be extracted from a sequence, such as the one shown in Figure 2.1. Once a call is extracted, the chirp type and parameters need to be determined. Call analysis is complicated by the large number of parameters associated with such methods – start position of a call, its duration, chirp type, and chirp parameters. Additionally, it can be difficult to incorporate the amplitude modulation of calls within the model due to the complicated and varied envelopes involved.

The solution considered here is to produce a time-frequency representation by sequentially estimating the instantaneous frequency of the recorded sequence. Using this time-frequency representation, individual calls can be extracted easily and further analysis can be performed.

To start with, a sinusoidal model is chosen for the estimation of the instantaneous frequency of the signal. Since a sinusoidal model is being used, the model will be unable to account for impulsive, or broadband components in signals. However, as seen in the STFTs of calls in the previous chapters, bat echolocation calls consist predominantly of modulated tones thus permitting the use of such a model.

The following section will introduce the basic sinusoidal model first. The basic model is prone to certain limitations. In order to overcome these limitations, the model can be extended, thus improving its robustness. The limitations and extensions will subsequently be discussed in Section 4.3.

4.2 A Basic Sinusoidal Model

The basic sinusoidal model consists of modelling the observed signal as a sum of weighted sinusoids [17]. Each sinusoid corresponds to a single tone present in the signal. Associated with each tone is a frequency, amplitude and phase, three parameters which completely characterise the tone. The estimation problem can be described as one where the number of tones and their parameters need to be estimated for the given observed signal. The sinusoidal model can be

written as [17, 62]:

$$x_t = \sum_{p=1}^{P} \left(a^{(p)} \cos(2\pi f^{(p)} t) + b^{(p)} \sin(2\pi f^{(p)} t) \right) + e_t \tag{4.1}$$

where x_t is the observation at time t, P is the number of tones present in the signal, $(a^{(p)}, b^{(p)})$ are the magnitudes of the cosine and sine components of the p^{th} tone and e_t represents a noise term. The formulation used here is equivalent to one written in the form of a magnitude and phase of a tone since the magnitude of the signal is obtained as $\sqrt{(a^{(p)})^2 + (b^{(p)})^2}$ and the phase as $\arctan\left(\frac{b^{(p)}}{a^{(p)}}\right)$.

In the method adopted here, the above sinusoidal model is used to evaluate the frequency content over a small segment of the observed signal. Each such segment, obtained using a sliding window, is termed a *frame*. By sliding a window over the observed signal and estimating the frequency content within each frame, it is possible to obtain a sequential estimate of the instantaneous frequency of the signal. The implicit assumption here is that the signal satisfies the model within the window, *i.e.*, within the window, the signal can be decomposed as a set of unmodulated tones.

Each frame can be treated independently of the others and consequently, the model specified in equation (4.1) needs to be specified accordingly. The length of the k^{th} frame is L and the centre of the frame occurs at time $t_k \in \tau = \{0, \ldots, T-1\}$, where T is the total length of the signal. The frequency content of the k^{th} frame is evaluated over the segment $l_k \in \{t_k - L', \ldots, t_k + L'\}$, where $L' = (L-1)/2$, L odd. The frame subscript, k, is used to indicate the parameters of the k^{th} frame. If P_k frequency components are present in the k^{th} frame, the observation can be written as:

$$x_{l_k} = \sum_{p=1}^{P_k} \left(a_k^{(p)} \cos \phi_{k,l_k - t_k}^{(p)} + b_k^{(p)} \sin \phi_{k,l_k - t_k}^{(p)} \right) + e_{l_k} \tag{4.2a}$$

where

$$\phi_{k,l}^{(p)} = 2\pi f_k^{(p)} l. \tag{4.2b}$$

In equations (4.2a) and (4.2b), x_{l_k} is the observation sample at time l_k and $e_{l_k} \sim \mathcal{N}(0, \sigma_{(e),k}^2)$ is zero-mean white Gaussian observation noise with variance $\sigma_{(e),k}^2$. The time-varying parameters of the model are given by $\left(a_k^{(p)}, b_k^{(p)}, f_k^{(p)}\right)$ which represent the amplitudes of the cosine and sine contributions, and the frequency of the p^{th} component respectively.

Equation (4.2a) and l_k are specified such that the analysis is centred within the frame. It should also be noted that in equation (4.2a), ϕ is *not* simply a function of the time instant t_k, but a function of position within the frame, as indicated by $l_k - t_k$. Figure 4.1 illustrates the decomposition of the signal within the frame onto a sinusoidal basis in the time-frequency domain. The figure is zoomed in over the time frame 0.4-0.6 s. The figure shows that the model provides a poor approximation to the frequency modulated signals, and extensions to the model will be considered in the following section to reduce this mismatch.

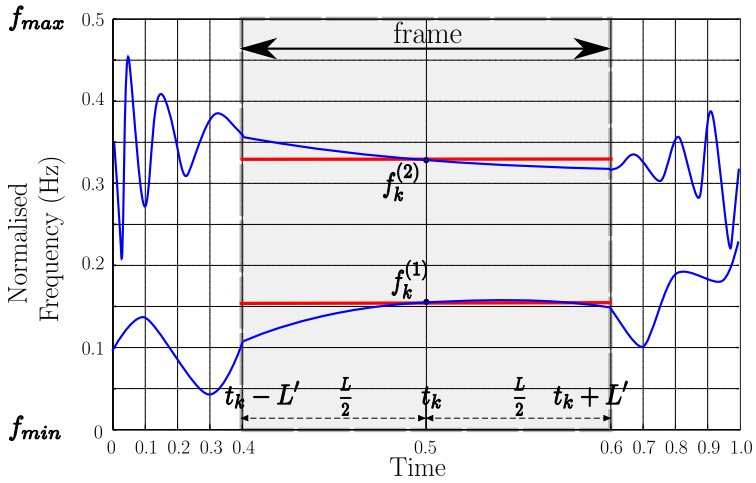


Figure 4.1: *A close up of the time-frequency representation of a frame and its context in the signal is shown. Within the frame, the signal is modelled as a sum of unmodulated tones.*

To allow for frequency and amplitude modulation of the signal over time, a second model is required to describe the evolution of the parameters, *i.e.*, the number of tones, their frequencies and amplitudes, over time. For this second model, a random walk can be chosen on the parameters [70]. The update equations for the parameters of the p^{th} tone are written as:

$$\begin{aligned} t_{k+1} &= t_k + \Delta t \\ P_{k+1} &= P_k + v_{(P),k} \end{aligned} \tag{4.3a}$$

$$\begin{aligned} f_{k+1}^{(p)} &= f_k^{(p)} + v_{(f),k}^{(p)} \\ a_{k+1}^{(p)} &= a_k^{(p)} + v_{(a),k}^{(p)} \\ b_{k+1}^{(p)} &= b_k^{(p)} + v_{(b),k}^{(p)} \end{aligned} \tag{4.3b}$$

where the subscript k indicates the value at the k^{th} frame, Δt determines the amount of overlap of subsequent frames, $v_{(P),k}$ is the process noise which determines whether the number of tones increases or decreases, $\left(v^{(p)}_{(f),k}, v^{(p)}_{(a),k}, v^{(p)}_{(b),k}\right)$ are respectively the process update noise for the model parameters f, a, and b. For the update of the number of tones P_k, an update matrix can be specified which increments or decrements the number of tones, or leaves the number unaltered at each update. The process noise for update of the tone parameters can be specified as a random walk such that:

$$\begin{aligned} v^{(p)}_{(f),k} &\sim \mathcal{N}\left(0, \sigma^2_{(f),k}\right) \\ v^{(p)}_{(a),k} &\sim \mathcal{N}\left(0, \sigma^2_{(a),k}\right) \\ v^{(p)}_{(b),k} &\sim \mathcal{N}\left(0, \sigma^2_{(b),k}\right) \end{aligned} \tag{4.4}$$

where $\sigma^2_{(f),k}$, $\sigma^2_{(a),k}$, $\sigma^2_{(b),k}$ are the variances of the distributions on the frequency and amplitude parameters respectively.

The term Δt in equation (4.3a) takes on a predefined value and is not estimated. Smaller values of Δt will produce more closely spaced estimates while large values will produce estimates farther apart due to the reduced overlap between subsequent frames.

4.3 Extensions to the Basic Sinusoidal Model

The harmonic model described above has certain limitations. The signals being analysed here, *viz.*, bat echolocation calls, are both frequency and amplitude modulated. The length of this window governs the variance on the resulting frequency estimates. The Cramer-Rao lower bound (CRLB) provides the lowest variance on the unknown parameters which may be attained by an estimator. The CRLB has been derived for the estimation of a single constant amplitude polynomial phase (or chirp) signal [17, 99, 100] illustrating the best achievable accuracy. It is seen that the variance on the frequency estimates has an inverse dependence on the window length. It then becomes possible to increase the accuracy of the estimates by increasing the length of the window thereby making use of more data.

In the model used here, the parameters are estimated within a sliding window of length L. The variance on the resulting frequency estimates then depends on the length of the window, and when L is small, the variance on the parameters will be larger than when L is large. Conversely, using a larger window may cause the model assumptions to be invalidated. For example, within

a long window, the windowed signal may be both frequency as well as amplitude modulated. While the preference may be to use a longer window thus improving the accuracy of the estimates, the limitations of the model constrain the extent to which the window length can be increased.

The model described in equation (4.2a) also does not explicitly account for harmonics which may be present in the signal and any harmonics will be estimated as independent tones. However, due to the presence of harmonics in echolocation calls, estimation of higher harmonics is useful in the model. Models which account for harmonics are commonly used for the analysis of musical signals [70, 101] and are well established. The following sections examine these harmonic models as well as extensions to the basic model to improve the flexibility and robustness of the model, particularly with regards to the use of longer window lengths.

4.3.1 Dealing with Harmonics

In some signal analysis problems, the signal of interest contains harmonics or overtones. For example, speech and musical signals commonly contain several harmonics, and in fact, this is also true of bat echolocation calls. Using the model described in equation (4.2a), each of these harmonics would be modelled as a separate tone.

In such cases where the signal contains harmonics, it can be more advantageous to estimate the frequency of the fundamental and the number of harmonics rather than the frequency of each harmonic component. The advantage stems from the need to estimate two parameters in the former case, a single frequency and the number of harmonics, as opposed to multiple *independent* frequencies.

The basic sinusoidal model is extended such that $f_k^{(p)}$ now represents the frequency of the p^{th} *fundamental tone* and all harmonics, or overtones, are obtained as integer multiples of the fundamental tone. Rewriting the observation model in terms of a fundamental tone and additional harmonics gives:

$$x_{l_k} = \sum_{p=1}^{P_k} \sum_{h=1}^{H_k^{(p)}} \left(a_k^{(p,h)} \cos \phi_{k,l_k-t_k}^{(p,h)} + b_k^{(p,h)} \sin \phi_{k,l_k-t_k}^{(p,h)} \right) + e_{l_k} \tag{4.5a}$$

where

$$\phi_{k,l}^{(p,h)} = 2\pi h f_k^{(p)} l. \tag{4.5b}$$

In equation (4.5a), $H_k^{(p)}$ is the number of overtones associated with the p^{th} tone. The amplitudes of the cosine and sine contributions of the h^{th} harmonic of the p^{th} tone are respectively given by $a_k^{(p,h)}$ and $b_k^{(p,h)}$. In contrast with the basic model described earlier in Section 4.2, $\phi_{k,l}^{(p,h)}$ is now a function of h as well. Comparisons between this model and the basic model will be considered in Chapter 6 to illustrate the advantages of accounting for additional harmonics of the signal.

When dealing with certain musical signals, it is seen that higher harmonics deviate from integer multiples of the fundamental tone [38]. The presence of deviation of harmonics in echolocation calls has not been previously considered, and one of the aims of this work is to establish the absence or presence of any such overtone deviation in the calls. The overtone deviation can be incorporated into the model by specifying an overtone deviation parameter μ, and the term $\phi_{k,l}^{(p,h)}$ can then be redefined as [38]:

$$\phi_{k,l}^{(p,h)} = 2\pi h f_k^{(p)}(1 + \mu_k^{(p,h)})l \tag{4.6}$$

where $\mu_k^{(p,h)}$ indicates deviation of the h^{th} harmonic from an integer multiple of the fundamental tone $f_k^{(p)}$. An additional constraint is added so that $\mu_k^{(p,1)} = 0$. In the model, $\mu_k^{(p,h)}$ is treated as a time-varying parameter. If $\mu_k^{(p,h)}$ is in fact a static parameter, an alternative approach may required which takes this into account (see, for example [95, 102] where a particle filter is used to estimate time-varying parameters while explicitly accounting for static parameters).

Similar to equations (4.3a) and (4.3b), the update for the number of harmonics and the harmonic deviation can be written as:

$$\begin{aligned} H_{k+1}^{(p)} &= H_k^{(p)} + v_{(H),k}^{(p)} \\ \mu_{k+1}^{(p,h)} &= \mu_k^{(p,h)} + v_{(\mu),k}^{(p,h)} \end{aligned} \tag{4.7}$$

where $v_{(H),k}^{(p)}$ is the process noise which determines how the number of harmonics will change, and $v_{(\mu),k}^{(p,h)} \sim \mathcal{N}(0, \sigma_{(\mu),k}^2)$.

When the number of harmonics is unknown, it is possible to over-model the system by assuming a constant and large number of harmonics in the model as is done by Dubois and Davy [70], and described in Section 3.5. Once the parameters of all tones and harmonics in the model are estimated, true harmonics can be detected by examining their associated magnitudes, *i.e.*, harmonics which are not present will have very low magnitudes associated with them. Thus, it

is possible to retain only valid harmonics by thresholding the magnitudes [70].

The advantage of such a method is that it is possible to easily detect a signal with only even or odd harmonics since thresholding is performed on the estimated magnitude. The downside to this approach, however, is the need to identify the precise value of the threshold since too large a threshold will result in ignoring weak harmonics, while too low a threshold will result in spurious harmonics.

An alternative method is to determine the number of harmonics at each instant. This can be performed using a RJMCMC sampler [38], and this is the approach adopted in the development of a sequential frequency estimation algorithm in Chapter 6.

4.3.2 Dealing with Signal Modulation

When dealing with long windows, the harmonic model may be unable to correctly model the observed signal due to the presence of frequency and/or amplitude modulation which occurs within the window. As discussed earlier, the motivation for using a longer window is to reduce the variance on the parameter estimates. A method to overcome this limitation is to then account for such modulation schemes within the harmonic model.

In particular, it is possible to account for the presence of some frequency modulation of tones within the window in the model. This can be done by modelling the signal within the window as a frequency modulated tone. To this end, the following sections will first introduce a linear chirp basis and subsequently, a general polynomial chirp basis.

While the frequency modulation is explicitly modelled within the window, it is chosen to not model the amplitude modulation within the window in this work. In particular, if the length of the window is not made arbitrarily long, it can be assumed that the observed signal envelope within the window is constant and the amplitude modulation can be ignored, thus preventing further complications to the model. This assumption may not always hold in practice and can lead to erroneous estimates when it is violated. The effect of this assumption and methods to overcome the limitations are considered in Chapter 7.

4.3.2.1 A Linear Chirp Basis

Frequency modulation within the sliding window can be characterised to some extent through a simple extension of the sinusoidal basis in equation (4.2b) to a linear chirp basis. Within the sliding window, the frequency at time t_k, at the centre of the window, is specified as $f_k^{(p)}$. Additionally, each fundamental tone is also described by a chirp rate parameter c_k.

Figure 4.2 shows an example of how a linear chirp basis is used to estimate the instantaneous frequencies of the signal components within a frame. By adopting a linear chirp rather than a sinusoidal basis, the mismatch between the windowed observations and the model is reduced, and a closer approximation is achieved. Incorporating the linear chirp basis into the model requires a redefinition of $\phi_{k,l}^{(p)}$ in equation (4.2b):

$$\phi_{k,l}^{(p)} = 2\pi \left(f_k^{(p)} + \frac{c_k^{(p)}}{2} l \right) l \tag{4.8}$$

Using the above definition of $\phi_{k,l}^{(p)}$, the windowed signal is modelled as a linear chirp within the window. Additionally, the analysis is centred within the window, which is to say that the instantaneous frequency of the p^{th} tone at the centre of the window is specified by the parameter $f_k^{(p)}$.

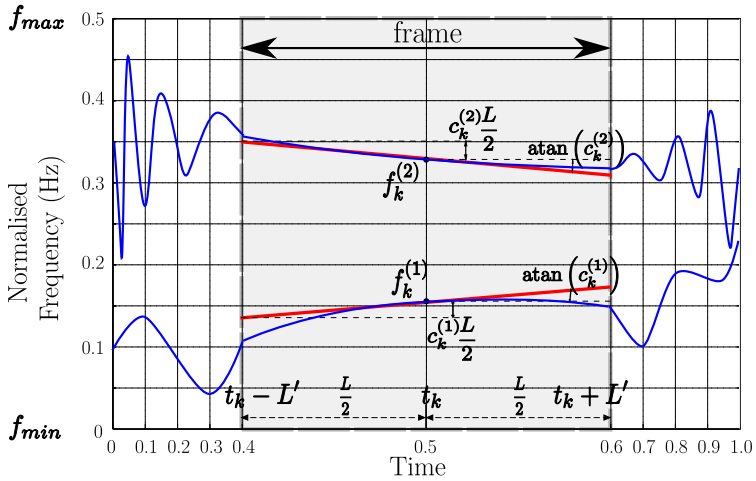


Figure 4.2: *A close up of the time-frequency representation of a frame and its context in the signal is shown. Within the frame, the signal is modelled as a sum of linear chirps which are shown as tangents to the frequency tracks. In this example, there are two linear chirps with instantaneous frequency and chirp-rate parameter pairs* $(f_k^{(1)}, c_k^{(1)})$, $(f_k^{(2)}, c_k^{(2)})$.

The linear chirp basis can easily be combined with the model specified in the previous section to take into account harmonics of the tone as:

$$\phi_{k,l}^{(p,h)} = 2\pi h \left((f_k^{(p)} + \frac{c_k^{(p)}}{2} l)(1 + \mu_k^{(p,h)}) \right) l \tag{4.9}$$

with similar extensions to higher order models as considered in the next section. In the above equation, both the chirp rate and overtone deviation parameters may be considered small such that their product becomes negligible. Consequently, the model is simplified to yield:

$$\phi_{k,l}^{(p,h)} = 2\pi h \left(f_k^{(p)}(1 + \mu_k^{(p,h)}) + \frac{c_k^{(p)}}{2} l \right) l \tag{4.10}$$

where the overtone deviation parameter μ only applies to the frequency term. The parameter state update equations will reflect the addition of the chirp rate parameter:

$$\begin{array}{rcccl} f_{k+1}^{(p)} & = & f_k^{(p)} + c_k^{(p)} \Delta t & + & v_{(f),k}^{(p)} \\ c_{k+1}^{(p)} & = & c_k^{(p)} & + & v_{(c),k}^{(p)} \end{array} \tag{4.11}$$

where $v_{(c),k}^{(p)} \sim \mathcal{N}(0, \sigma_{(c),k}^2)$ is the process noise for the chirp rate update.

4.3.2.2 A General Polynomial Chirp Basis

The linear chirp basis specified above provides more flexibility within the model than a pure sinusoidal basis. A further redefinition of $\phi_{k,l}^{(p,h)}$ allows the model to be extended to comprise a general n^{th} order polynomial basis. This, the most general form of the model, will be referred to as the general harmonic model (GHM).

A general n^{th} order polynomial chirp can be used to describe the instantaneous frequency $f(l_k)$, at any point l_k within the sliding window as:

$$f(l_k) = s_{0,k} + s_{1,k}\, l_k + \frac{s_{2,k}\, l_k^2}{2!} + \cdots + \frac{s_{n,k}\, l_k^n}{n!} = \sum_{j=0}^{n} \frac{s_{j,k}\, l_k^j}{j!} \tag{4.12}$$

and since the instantaneous frequency is obtained as the derivative of the phase, $f_t = \dfrac{1}{2\pi} \dfrac{d\phi}{dt}$,

the instantaneous phase is given by:

$$\phi(l_k) = 2\pi \left(s_{0,k} + \frac{s_{1,k} l_k}{2!} + \frac{s_2 l_k^2}{3!} + \cdots + \frac{s_n l_k^n}{(n+1)!} \right) l_k + \phi_0 \tag{4.13}$$

where ϕ_0 represents some initial phase and is assumed to be zero.

Using equation (4.13), and incorporating additional overtones as discussed in Section 4.3.1, the observation model is rewritten as:

$$x_{l_k} = \sum_{p=1}^{P_k} \sum_{h=1}^{H_k^{(p)}} \left(a_k^{(p,h)} \cos \phi_{k,l_k - t_k}^{(p,h)} + b_k^{(p,h)} \sin \phi_{k,l_k - t_k}^{(p,h)} \right) + e_{l_k} \tag{4.14}$$

where

$$\phi_{k,l}^{(p,h)} = 2\pi h \left(s_{0,k}^{(p)} (1 + \mu_k^{(p,h)}) + \frac{s_{1,k}^{(p)} l}{2!} + \cdots + \frac{s_{n,k}^{(p)} l^n}{(n+1)!} \right) l \tag{4.15}$$

Using a first-order polynomial, $s_{0,k}$ corresponds to the fundamental frequency term f_k and $s_{1,k}$ corresponds to the chirp-rate c_k. Additionally, in equation (4.15), the overtone deviation term μ affects only the fundamental frequency term $s_{0,k}$.

While a larger polynomial order allows characterisation of greater detail within the sliding window, increasing the order also requires estimation of additional parameters . There is thus a trade-off between model simplicity, governed by the polynomial order, and the extent of frequency modulation which can be modelled within the sliding window.

The parameter state update equations for the general polynomial basis can be written as follows, with the subscript (p) dropped for clarity:

$$\begin{aligned}
s_{0,k+1} &= s_{0,k} + s_{1,k}\Delta t + s_{2,k}\frac{(\Delta t)^2}{2!} + \ldots + s_{n,k}\frac{(\Delta t)^n}{n!} &+ v_{(s_0),k} \\
s_{1,k+1} &= s_{1,k} + s_{2,k}\Delta t + s_{3,k}\frac{(\Delta t)^2}{2!} + \ldots + s_{n,k}\frac{(\Delta t)^{(n-1)}}{(n-1)!} &+ v_{(s_1),k} \\
&\vdots \qquad\qquad\qquad\qquad \vdots &\qquad \vdots \\
s_{n,k+1} &= s_{n,k} &+ v_{(s_n),k}
\end{aligned} \tag{4.16}$$

or more compactly as:

$$s_{i,k+1} = \sum_{j=i}^{n} s_{j,k} \frac{(\Delta t)^{(j-i)}}{(j-i)!} + v_{(s_i),k} \tag{4.17}$$

where $\{v_{(s_0),k}, v_{(s_1),k}, \ldots, v_{(s_n),k}\}$ are the process noise terms for the update of the chirp pa-

rameters such that $v_{(s_n),k} \sim \mathcal{N}(0, \sigma^2_{(s_n),k})$.

The above set of equations can be written more compactly in terms of matrices. The vector of parameters is given by $\mathbf{s}_k^{(p)} = [s_{0,k}^{(p)}, \ldots, s_{n,k}^{(p)}]^T$ and the vector of noise terms by $\mathbf{v}_{(s),k}^{(p)} = [v_{(s_0),k}^{(p)}, \ldots, v_{(s_n),k}^{(p)}]^T$. The update equation is then obtained as:

$$\mathbf{s}_{k+1}^{(p)} = \Delta_s \mathbf{s}_k^{(p)} + \mathbf{v}_{(s),k}^{(p)} \tag{4.18}$$

where the matrix Δ_s is defined as:

$$\Delta_s = \begin{bmatrix} 1 & \Delta t & \frac{(\Delta t)^2}{2!} & \frac{(\Delta t)^3}{3!} & \cdots & \frac{(\Delta t)^n}{n!} \\ 0 & 1 & \Delta t & \frac{(\Delta t)^2}{2!} & \cdots & \frac{(\Delta t)^{(n-1)}}{(n-1)!} \\ 0 & 0 & 1 & \Delta t & \cdots & \frac{(\Delta t)^{(n-2)}}{(n-2)!} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{4.19}$$

The choice of polynomial order n is dependent on the signal of interest. Chapter 7 will compare the performance of different order polynomials and their relevance to the analysis of bat echolocation calls.

4.4 Frequency Estimation and the General Linear Model

The general linear model (GLM) [17, 62] provides a useful generic method for describing a large class of models. In [62], the authors state that any data which can be represented in terms of a "linear combination of basis functions with an additive Gaussian noise component satisfies the general linear model". More formally, for a sequence of observations of length K, the observation x_k at time k can be written as:

$$x_k = \sum_{p=1}^{P} a^{(p)} g_k^{(p)} + e_k \tag{4.20}$$

where $g_k^{(p)}$ constitutes the basis, $a^{(p)}$ determines the contribution of that basis and e_k is the observation noise.

The observation equation can be written in matrix-vector form as:

$$\mathbf{x} = \mathbf{G}\mathbf{a} + \mathbf{e} \tag{4.21}$$

where $\mathbf{x} = [x_1, \ldots, x_K]^T$, $\mathbf{a} = [a^{(1)}, \ldots, a^{(P)}]^T$, $\mathbf{e} = [e_1, \ldots, e_K]^T$ and the matrix G is defined as:

$$\mathbf{G} = \begin{bmatrix} g_1^{(1)} & g_1^{(2)} & \cdots & g_1^{(P)} \\ g_2^{(1)} & g_2^{(2)} & \cdots & g_2^{(P)} \\ \vdots & \vdots & \ddots & \vdots \\ g_K^{(1)} & g_K^{(2)} & \cdots & g_K^{(P)} \end{bmatrix} \tag{4.22}$$

The following section demonstrates how the harmonic model can be placed within the GLM framework before subsequently examining the likelihood function and prior distributions.

4.4.1 Frequency Estimation Context

The GHM described in equations (4.14) and (4.15) can be written in terms of the GLM. The cosine and sine components correspond to the basis functions in the GLM and together with the amplitude parameters, the observation is written as a linear combination of cosine and sine components.

Due to the time-varying nature of the GHM, the basis functions and their weights are time-varying. The subscript k indicates the parameters within a frame. The observation vector $\mathbf{x}_k$ is defined as $\mathbf{x}_k = [x_{t_k-L'}, \ldots, x_{t_k+L'}]^T$ and $\mathbf{e}_k = [e_{t_k-L'}, \ldots, e_{t_k+L'}]^T$ is the white Gaussian noise vector.

The vector of harmonics associated with the tones is given by $\mathbf{H}_k = [H_k^{(1)}, \ldots, H_k^{(P_k)}]^T$ and the total number of components (tones and their harmonics) is $O_k = \sum_{p=1}^{P_k} H_k^{(p)}$. The vector of cosine and sine amplitudes of the p^{th} tone (including the harmonics) is specified as:

$$A_k^{(p)} = [a_k^{(p,1)}, \ldots, a_k^{(p,H_k^{(p)})}, b_k^{(p,1)}, \ldots, b_k^{(p,H_k^{(p)})}]^T \tag{4.23a}$$

such that

$$\mathbf{a}_k = [{A_k^{(1)}}^T, \ldots, {A_k^{(P_k)}}^T]^T \tag{4.23b}$$

is the vector of all amplitudes.

Depending on the choice of model, the parameters which define the frequency modulation structure of each tone within the window may be a) frequency, b) frequency and chirp rate, or c) an n^{th} order polynomial. The first two options are merely specific cases where the polynomial order is restricted to zero and one respectively. However, since these particular options will be used more extensively in the following chapters, these particular instances of the general polynomial will be explicitly specified here.

When using a general n^{th} order polynomial as discussed in Section 4.3.2.2, the parameters of the p^{th} tone is given by $\mathbf{s}_k^{(p)} = [s_{0,k}^{(p)}, \ldots, s_{n,k}^{(p)}]^T$, such that $\mathbf{S}_k = [{\mathbf{s}_k^{(1)}}^T, \ldots, {\mathbf{s}_k^{(P_k)}}^T]^T$. Constraining the model to a zeroth-order polynomial results in $f_k^{(p)} = s_{0,k}^{(p)}$ and in the case of a first order polynomial, additionally, $c_k^{(p)} = s_{1,k}^{(p)}$. In these cases, the vector of frequencies is (alternatively) specified as $\mathbf{f}_k = [f_k^{(1)}, \ldots, f_k^{(P_k)}]^T$, and the vector of chirp rates as $\mathbf{c}_k = [c_k^{(1)}, \ldots, c_k^{(P_k)}]^T$.

When overtone deviation is considered, the deviation of harmonics of the p^{th} tone is indicated by the vector $M_k^{(p)} = [\mu_k^{(p,1)}, \ldots, \mu_k^{(p,H_k^{(p)})}]^T$, and $\mathbf{M}_k = [{M_k^{(1)}}^T, \ldots, {M_k^{(P_k)}}^T]^T$.

The matrix of basis functions $\mathbf{G}_k$ is constructed using the (modulated) tones specified by the parameter vector $\mathbf{S}_k$. The model can be written in matrix-vector form by defining the following matrices:

$$\mathbf{\Phi}_k = \begin{bmatrix} \phi_{k,-L'}^{(1,1)} & \cdots & \phi_{k,-L'}^{(1,H_k^{(1)})} & \phi_{k,-L'}^{(2,1)} & \cdots & \phi_{k,-L'}^{(P_k,H_k^{(P_k)})} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \phi_{k,L'}^{(1,1)} & \cdots & \phi_{k,L'}^{(1,H_k^{(1)})} & \phi_{k,L'}^{(2,1)} & \cdots & \phi_{k,L'}^{(P_k,H_k^{(P_k)})} \end{bmatrix} \tag{4.24a}$$

where the general definition of $\phi_{k,l}^{(p,h)}$ is described in equation (4.15). The $\mathbf{\Phi}_k$ matrix specifies the instantaneous phase of the (modulated) tones which form the basis of the $\mathbf{G}_k$ matrix which is then obtained as:

$$\mathbf{G}_k = [\cos(\mathbf{\Phi}_k) \quad \sin(\mathbf{\Phi}_k)] \tag{4.24b}$$

where $\cos(\cdot)$ and $\sin(\cdot)$ operate element-wise on $\mathbf{\Phi}_k$ to produce the augmented $\mathbf{G}_k$ matrix with dimensions $L \times 2O_k$. Using the definitions of $\mathbf{a}_k$ and $\mathbf{G}_k$ in equations (4.23b) and (4.24b), the harmonic model can be written in the framework of the GLM.

4.4.2 The Likelihood Function and Prior Distributions

This section describes the likelihood function for the GHM described previously, as well as the prior distributions used in implementing the Bayesian solution.

Since the observation noise is Gaussian, if $\boldsymbol{\psi}_k = \{P_k, \mathbf{H}_k, \mathbf{S}_k, \mathbf{M}_k, \mathbf{a}_k, \sigma^2_{(e),k}\}$, the likelihood function is written as:

$$p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) = \frac{1}{(2\pi\sigma^2_{(e),k})^{\frac{L}{2}}} \exp\left[-\frac{\|\mathbf{x}_k - \mathbf{G}_k\mathbf{a}_k\|^2}{2\sigma^2_{(e),k}}\right]. \tag{4.25}$$

Since a Bayesian method is adopted in estimating the frequencies present in the signal, it is necessary to specify prior distributions for the parameters. The likelihood function itself does not draw any distinction between harmonics and fundamental frequencies that have the same frequency, *i.e.*, a harmonic can be replaced by a fundamental tone of the same frequency with no impact on the likelihood function. The likelihood function is merely a goodness of fit test and penalty terms arising from a large number of tones and harmonics need to be incorporated via the prior distributions. A set of uninformative priors are used in [29, 33, 38] for the estimation of sinusoids as well as general frequency modulated signals and these priors are used here as well.

For the number of tones P_k, a truncated Poisson distribution is selected

$$p(P_k \mid \lambda_t) = \frac{\lambda_t^{P_k} \exp(\lambda_t)}{P_k!}, P \in [P_{min}, P_{max}] \tag{4.26}$$

where λ_t may be considered as the number of expected frequency components in the signal [29]. As described in [38], a truncated Poisson distribution is specified for the prior on the number of harmonics of each tone as

$$p(H_k^{(p)} \mid \lambda_h) = \frac{\lambda_h^{H_k^{(p)}} \exp(\lambda_h)}{H_k^{(p)}!}, H_k^{(p)} \in [H_{min}, H_{max}] \tag{4.27}$$

where λ_h indicates the number of expected harmonics.

A uniform distribution is chosen for each of the polynomial basis chirp parameters

$$p(s_{n,k}) = \mathcal{U}[s_{n,min}, s_{n,max}] \tag{4.28}$$

where $\mathcal{U}[a, b]$ indicates a uniform distribution over the limits $[a, b]$. This prior distribution is based on the assumption that the signal under consideration has been filtered through a lowpass filter and sampled at a sufficiently high sampling rate such that it does not possess any aliasing. Under the lowpass assumption, a suitable limit for the uniform distribution on the frequency is $\left[0, \frac{f_s}{2}\right]$, where f_s is the sampling frequency. Ranges for higher order terms in the polynomial may be dependent on the signal of interest, however, the fundamental limit imposed on the frequency, $\left[0, \frac{f_s}{2}\right]$, implicitly limits the range of the higher order terms since the frequency as given by equation (4.12) may not exceed the limits. Additionally, a uniform distribution is specified for the overtone deviation parameter [38]:

$$\mu_k^{(p)} \sim \mathcal{U}[\mu_{min}, \mu_{max}] \tag{4.29}$$

The noise variance is assumed to follow an inverse-gamma distribution, $\sigma^2_{(e),k} \sim \mathfrak{IG}(\alpha_e, \beta_e)$, with scale and shape parameters (α_e, β_e) respectively [29]. In other words:

$$p(\sigma^2_{(e),k} \mid \alpha_e, \beta_e) = \frac{\alpha_e^{\beta_e}}{\Gamma(\beta_e)}(\sigma^2_{(e),k})^{-(\beta_e+1)} \exp\left(-\frac{\alpha_e}{\sigma^2_{(e),k}}\right) \tag{4.30}$$

For the amplitude parameters, a multivariate normal distribution is defined over the $2O_k$ amplitude parameters, $\mathbf{a}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_k)$, $\mathbf{\Sigma}_k = \sigma^2_{(e),k}\delta_k^2\mathbb{I}_{2O_k}$, where $\mathbb{I}_Q$ is the $Q \times Q$ identity matrix.

$$p(\mathbf{a}_k \mid P_k, \mathbf{H}_k, \sigma^2_{(e),k}, \delta_k^2) = \frac{1}{(2\pi\delta_k^2\sigma^2_{(e),k})^{O_k}} \exp\left(-\frac{1}{2}\mathbf{a}_k^T\Sigma_k^{-1}\mathbf{a}_k\right) \tag{4.31}$$

An extra parameter, δ_k^2, is introduced from the prior on the amplitudes. The introduction of this parameter facilitates marginalisation of parameters from the posterior distribution which is considered in Section 4.4.4. The δ_k^2 parameter may be considered an indication of the expected SNR of the signal [29] and an inverse gamma prior is chosen $p(\delta_k^2) = \mathfrak{IG}(\alpha_\delta, \beta_\delta)$:

$$p(\delta_k^2 \mid \alpha_\delta, \beta_\delta) = \frac{\alpha_\delta^{\beta_\delta}}{\Gamma(\beta_\delta)}(\delta_k^2)^{-(\beta_\delta+1)} \exp\left(-\frac{\alpha_\delta}{\delta_k^2}\right). \tag{4.32}$$

4.4.3 Hyper-parameters

The prior distributions on the parameters have introduced several hyper-parameters: λ_t, λ_h, (α_e, β_e), $(\alpha_\delta, \beta_\delta)$. These hyper-parameters provide a means for incorporating prior knowl-

edge of the parameters. However, since the priors are uninformative, the exact values of the hyper-parameters are unable to adversely impact the posterior distribution even when the prior information is misleading [33].

For an inverse gamma distribution, setting $\alpha = \beta = 0$ reduces the prior to Jeffrey's prior [29,33] reflecting a lack of knowledge in the prior. Additionally, setting $\beta = 2$ in the distribution results in a uninformative prior since the distribution has infinite variance [33]. The precise value of α in this case is seen to not affect the estimation process significantly [29, 33]. Setting $\beta = 2$ ensures that the distribution can be sampled from while still ensuring that the prior is uninformative.

The number of expected tones and harmonics is indicated by λ_t and λ_h respectively. The values of λ_t and λ_h penalise the posterior distribution when the number of tones and/or harmonics deviates from the expected number. It is possible to further parameterise the hyper-parameters of the prior distributions which would be equivalent to estimating the values of the hyper-parameters by describing priors and transition probabilities. Such an approach is adopted, for example, by Dubois *et al.* in [48] where the process noise variance varies over time. In the work presented in the following chapters, however, further parameterisation of the hyper-parameters is not considered, with the chosen priors being of a sufficiently uninformative nature.

In addition to the priors on the parameters being uninformative, the priors on the noise variance, $\sigma^2_{(e),k}$, and amplitudes, $\mathbf{a}_k$, are also conjugate priors. The use of the conjugate priors aids in the process of parameter marginalisation which will be addressed in the following section.

4.4.4 Parameter Marginalisation

In certain cases, some parameters in the model may be of little or no interest. For example, in the time-frequency analysis problem, the time-varying frequency estimates may be of greater interest than the estimate of the amplitudes. Under certain circumstances, it is possible to marginalise these unwanted parameters out of the posterior distribution, thus eliminating them from the estimation process and reducing the parameter state space dimensions. An advantage of marginalisation, or Rao-Blackwellisation, is that the resulting estimator has a lower variance [36]. In the GLM, it is possible to analytically marginalise the amplitude and noise variance parameters, $\mathbf{a}_k$ and $\sigma^2_{(e),k}$ through the use of conjugate priors [62].

The amplitude and noise variance parameters are termed nuisance parameters and denoted by

$\boldsymbol{\xi}_k = \{\mathbf{a}_k, \sigma^2_{(e),k}\}$. The remaining parameters to be estimated are $\boldsymbol{\psi}'_k = \boldsymbol{\psi}_k \setminus \boldsymbol{\xi}_k$. The integration of the posterior distribution of the GLM over the amplitude and noise variance terms is illustrated in [17] and [62] using a uniform prior for the amplitudes and Jeffreys prior for the noise variance. The marginalisation process here is carried out along similar lines, but using the priors defined in the previous section.

The aim is to obtain a likelihood function $p(\mathbf{x}_k \mid \boldsymbol{\psi}'_k)$ which is independent of the nuisance parameters. This *marginalised likelihood function* will be utilised in the recursive Bayesian estimator described in the next chapter. The marginalised likelihood function, is obtained as:

$$p(\mathbf{x}_k \mid \boldsymbol{\psi}'_k) = \int p(\mathbf{x}_k, \boldsymbol{\xi}_k \mid \boldsymbol{\psi}'_k) d\boldsymbol{\xi}_k = \int p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) p(\boldsymbol{\xi}_k) d\boldsymbol{\xi}_k \tag{4.33}$$

$$= \iint p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) p(\sigma^2_{(e),k} \mid \alpha_e, \beta_e) p(\mathbf{a}_k \mid P_k, \mathbf{H}_k, \sigma^2_{(e),k}, \delta^2_k) d\sigma^2_{(e),k} d\mathbf{a}_k \tag{4.34}$$

Expanding equation 4.34 in terms of the distributions gives:

$$p(\mathbf{x}_k \mid \boldsymbol{\psi}'_k) = \iint \left(\begin{array}{c} \frac{1}{\left(2\pi\sigma^2_{(e),k}\right)^{\frac{L}{2}}} \exp\left(-\frac{\mathbf{x}_k^T\mathbf{x}_k - 2\mathbf{x}_k^T\mathbf{G}_k\mathbf{a}_k + \mathbf{a}_k^T\mathbf{G}_k^T\mathbf{G}_k\mathbf{a}_k}{2\sigma^2_{(e),k}} \right) \\ \times \frac{\alpha_e^{\beta_e}}{\Gamma(\beta_e)} (\sigma^2_{(e),k})^{-(\beta_e+1)} \exp\left(-\frac{\alpha_e}{\sigma^2_{(e),k}} \right) \\ \frac{1}{\left(2\pi\delta_k^2\sigma^2_{(e),k}\right)^{O_k}} \exp\left(-\frac{1}{2}\mathbf{a}_k^T\Sigma_k^{-1}\mathbf{a}_k \right) \end{array} \right) d\sigma^2_{(e),k} d\mathbf{a}_k \tag{4.35}$$

$$= \iint \left(\begin{array}{c} \frac{\alpha_e^{\beta_e}}{\Gamma(\beta_e)(2\pi)^{(\frac{L}{2}+O_k)}(\delta_k^2)^{O_k}(\sigma^2_{(e),k})^{(\frac{L}{2}+O_k+\beta_e+1)}} \\ \times \exp\left(-\frac{\mathbf{x}_k^T\mathbf{x}_k + 2\alpha_e - 2\mathbf{x}_k^T\mathbf{G}_k\mathbf{a}_k + \mathbf{a}_k^T\mathbf{F}_k\mathbf{a}_k}{2\sigma^2_{(e),k}} \right) \end{array} \right) d\sigma^2_{(e),k} d\mathbf{a}_k \tag{4.36}$$

where $\mathbf{F}_k = (\mathbf{G}_k^T\mathbf{G}_k + \delta_k^{-2}\mathbb{I}_{2O_k})$.

The integral over $\mathbf{a}_k$ is performed first using the identity [103, page 248]:

$$\int_{\mathbb{R}^P} \exp\left[-\frac{1}{2}\left(\alpha + 2\mathbf{y}^T\boldsymbol{\Lambda} + \mathbf{y}^T\boldsymbol{\Gamma}\mathbf{y} \right) \right] d\mathbf{y} = \frac{(2\pi)^{\frac{P}{2}}}{\sqrt{\det(\boldsymbol{\Gamma})}} \exp\left[-\frac{1}{2}\left(\alpha - \boldsymbol{\Lambda}^T\boldsymbol{\Gamma}^{-1}\boldsymbol{\Lambda} \right) \right] \tag{4.37}$$

where the integral is performed over the P-dimensional real space. Setting $\mathbf{y} = \mathbf{a}_k$, $\boldsymbol{\Gamma} =$

$\sigma_{(e),k}^{-2}\mathbf{F}_k$, and $\mathbf{\Lambda} = -\sigma_{(e),k}^{-2}\mathbf{G}_k^T\mathbf{x}_k$, the integral over $\mathbf{a}_k$ simplifies to:

$$p(\mathbf{x}_k \mid \boldsymbol{\psi}'_k) = \int \left(\begin{array}{c} \dfrac{\alpha_e^{\beta_e}}{\Gamma(\beta_e)(2\pi)^{\frac{L}{2}}(\delta_k^2)^{O_k}(\sigma_{(e),k}^2)^{(\frac{L}{2}+\beta_e+1)}\sqrt{\det(\mathbf{F}_k)}} \\ \times \exp\left(-\dfrac{\mathbf{x}_k^T\mathbf{x}_k + 2\alpha_e - \mathbf{x}_k^T\mathbf{G}_k\mathbf{F}_k^{-1}\mathbf{G}_k^T\mathbf{x}_k}{2\sigma_{(e),k}^2}\right) \end{array} \right) d\sigma_{(e),k}^2 \tag{4.38}$$

Following this, the gamma integral can be used to integrate out the noise variance [62, page 16]:

$$\int_0^\infty (\sigma^2)^{-(\beta+1)} \exp\left(-\frac{\alpha}{\sigma^2}\right) d\sigma^2 = \frac{\Gamma(\beta)}{\alpha^\beta} \tag{4.39}$$

where the integral is carried out over the interval $[0, \infty]$ since the variance in non-negative.

Setting $\alpha = \dfrac{1}{2}(\mathbf{x}_k^T\mathbf{x}_k + 2\alpha_e - \mathbf{x}_k^T\mathbf{G}_k\mathbf{F}_k^{-1}\mathbf{G}_k^T\mathbf{x}_k)$, $\beta = \dfrac{L}{2} + \beta_e$ and $\sigma = \sigma_{(e),k}$ reduces the marginalised likelihood function to:

$$p(\mathbf{x}_k \mid \boldsymbol{\psi}'_k) = \frac{2^{\beta_e}\alpha_e^{\beta_e}\Gamma(\frac{L}{2}+\beta_e)}{\pi^{\frac{L}{2}}\Gamma(\beta_e)} \frac{\left(\mathbf{x}_k^T\mathbf{x}_k + 2\alpha_e - \mathbf{x}_k^T\mathbf{G}_k\mathbf{F}_k^{-1}\mathbf{G}_k^T\mathbf{x}_k\right)^{-\left(\frac{L}{2}+\beta_e\right)}}{\delta_k^{2O_k}\sqrt{\det(\mathbf{F}_k)}}. \tag{4.40}$$

Since the harmonic model can be placed in the GLM framework, it becomes possible to use this method to marginalise parameters from the distribution. The marginalisation allows implementation of a filter of reduced complexity which will be discussed in the following chapter.

4.5 Summary

This chapter introduced the signal model which will be used for sequential frequency estimation. The design of the signal model is motivated by the signals of interest in this work, *viz.*, bat echolocation calls.

A basic model is first presented for the signal model, where the signal is approximated as a sum of unmodulated tones within a sliding window. A sinusoidal basis may be inadequate when dealing with frequency modulated signals. This can arise if the frequency modulation of the signal within the window causes the model to approximate a single tone with several closely spaced tones. One solution to this is to use a window of short duration, however, the use of a short window can lead to a comparatively large variance on the parameter estimates.

By way of extending and improving the model, the sinusoidal basis in the basic model is extended to a linear chirp basis, and subsequently, to a general n^{th} order polynomial chirp basis. These basis functions afford significantly more flexibility than the standard sinusoidal basis which is commonly used, and allow the use of longer windows due to the decreased mismatch between the observations and the signal model within the window.

Subsequent extensions to the basic model are considered by way of accounting for additional harmonics as well as deviations of the harmonics from integer multiples of the fundamental frequency. The modelling of harmonics is particularly important with regard to music signals which sometimes exhibit inharmonicity.

Chapter 7 also examines the use of different polynomial basis functions with regards to bat echolocation calls. A higher order polynomial basis requires estimation of a larger number of parameters, and consequently, there is a trade-off between model flexibility and model simplicity. That chapter examines parameter estimation accuracy using different order polynomials under constrained conditions.

Section 4.4 illustrates how the frequency estimation problem can be placed in the context of the general linear model (GLM). Uninformative conjugate priors are described for the parameters to describe a complete lack of information regarding the parameters. The use of these conjugate distributions allows for analytical marginalisation of the amplitudes of the tones as well as the noise variance which gives rise to a marginalised likelihood function. This marginalisation process is not specific to the particular problem of frequency estimation, but is applicable to any model which can be placed within the framework of the GLM. The marginalised likelihood function derived here is used in the following chapter to perform sequential frequency estimation through an implementation of a recursive Bayesian filter.

Chapter 5
A New Marginalised Particle Filter for Sequential Frequency Estimation

Introduction

As discussed in the previous chapter, a sequential approach to time-varying frequency estimation has been adopted in this work. The sequential model developed previously is sufficiently flexible to account for significant variations in the signal, making it suitable for the analysis of bat echolocation calls.

As introduced in Chapter 3, particle filters offer a means for implementing a recursive Bayesian filter. Particle filters offer an attractive solution to the sequential frequency estimation problem on account of the non-linear nature of the observation equation and the multi-modal posterior distribution.

The Rao-Blackwellised particle filter (RBPF), reviewed in Chapter 3, can be used to perform parameter marginalisation in the particle filter framework. The RBPF is based on achieving a separation of the parameter space into a set of linear and non-linear parameters, with a particle filter being used to track the non-linear parameters, and the Kalman filter, to track the linear parameters. As seen at the end of that chapter, Dubois and Davy [70] utilise the RBPF framework in estimating the time-varying frequency content of a signal. The authors develop a flexible and complete framework for sequential estimation of a time-varying number of frequency components.

As an alternative to the RBPF discussed in Chapter 3, a new marginalisation technique is developed in this chapter giving rise to the marginalised particle filter (MPF). The sequential updates of the amplitude parameters as described by a random walk can be tracked using the Kalman filter in the RBPF. If, on the other hand, the sequential relation as described by the random walk is ignored, and the amplitudes at each instant are assumed independent of the previous value, then by using a set of conjugate priors, it is possible to integrate out the parameters. Using a similar assumption, the observation noise variance is also marginalised out from the posterior

distribution. This allows the marginalised likelihood function developed in the previous chapter to be used to estimate the instantaneous frequency of the signal.

It will be shown in Section 5.4 that the marginalised form of the particle filter derived here provides a good alternative option to the well established RBPF for application to the linear and Gaussian class of problems.

The following section develops this new form of the particle filter with subsequent sections discussing its application to the sequential frequency estimation problem. This new frequency estimation algorithm builds and improves upon established techniques for both sequential as well as batch offline frequency estimation.

5.1 Developing the Marginalised Particle Filter

In this section, an alternative approach to the standard RBPF is considered for parameter marginalisation. The set of parameters is partitioned into those which are of interest $\boldsymbol{\psi}'_k$, and those which may be termed *nuisance parameters* $\boldsymbol{\xi}_k$. The nuisance parameters $\boldsymbol{\xi}_k$, may either not be of interest, or it may be possible to estimate these conditional on $\boldsymbol{\psi}'_k$, for example, using a least squares estimate.

In order to marginalise the nuisance parameters $\boldsymbol{\xi}_k$ from the posterior distribution, $p(\boldsymbol{\Psi}_k \mid \mathbf{X}_k)$, it is necessary to integrate the posterior distribution with respect to $\boldsymbol{\Xi}_k = [\boldsymbol{\xi}_0, \ldots, \boldsymbol{\xi}_k]$. The posterior distribution satisfies the recursion [82]:

$$p(\boldsymbol{\Psi}_k \mid \mathbf{X}_k) = p(\boldsymbol{\Psi}_{k-1} \mid \mathbf{X}_{k-1}) \frac{p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) p(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_{k-1})}{p(\mathbf{x}_k \mid \mathbf{X}_{k-1})}. \tag{5.1}$$

The distribution over the reduced parameter set, $\boldsymbol{\Psi}'_k = \boldsymbol{\Psi}_k \setminus \boldsymbol{\Xi}_k$, is now given by:

$$p(\boldsymbol{\Psi}'_k \mid \mathbf{X}_k) = \int p(\boldsymbol{\Psi}_k \mid \mathbf{X}_k) d\boldsymbol{\Xi}_k \tag{5.2}$$

It is assumed that the parameters $\boldsymbol{\psi}'_k$ and $\boldsymbol{\xi}_k$ are independent of each other. This may not necessarily be true, however, this assumption is necessary to achieve a tractable marginalisation

of the nuisance parameters $\boldsymbol{\xi}_k$. The transition prior $p(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_{k-1})$ can then be written as:

$$\begin{aligned} p(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_{k-1}) &= p(\boldsymbol{\psi}'_k, \boldsymbol{\xi}_k \mid \boldsymbol{\psi}'_{k-1}, \boldsymbol{\xi}_{k-1}) && (5.3) \\ &= p(\boldsymbol{\psi}'_k \mid \boldsymbol{\psi}'_{k-1}, \boldsymbol{\xi}_k, \boldsymbol{\xi}_{k-1}) p(\boldsymbol{\xi}_k \mid \boldsymbol{\xi}_{k-1}, \boldsymbol{\psi}'_{k-1}) && (5.4) \\ &= p(\boldsymbol{\psi}'_k \mid \boldsymbol{\psi}'_{k-1}) p(\boldsymbol{\xi}_k \mid \boldsymbol{\xi}_{k-1}) && (5.5) \end{aligned}$$

Substituting equation (5.1) into equation (5.2) and rewriting the integral in terms of $\boldsymbol{\psi}'_k$ and $\boldsymbol{\xi_k}$ then yields:

$$\begin{aligned} p(\boldsymbol{\Psi}'_k \mid \mathbf{X}_k) &= \int p(\boldsymbol{\Psi}_{k-1} \mid \mathbf{X}_{k-1}) \frac{p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) p(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_{k-1})}{p(\mathbf{x}_k \mid \mathbf{X}_{k-1})} d\boldsymbol{\Xi}_k && (5.6) \\ &\propto \int p(\boldsymbol{\Psi}_{k-1} \mid \mathbf{X}_{k-1}) p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) p(\boldsymbol{\psi}'_k \mid \boldsymbol{\psi}'_{k-1}) p(\boldsymbol{\xi}_k \mid \boldsymbol{\xi}_{k-1}) d\boldsymbol{\Xi}_k && (5.7) \\ &\propto \int \left(\begin{array}{l} p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) p(\boldsymbol{\psi}'_k \mid \boldsymbol{\psi}'_{k-1}) p(\boldsymbol{\xi}_k \mid \boldsymbol{\xi}_{k-1}) \\ \times \cdots \times p(\mathbf{x}_1 \mid \boldsymbol{\psi}_1) p(\boldsymbol{\psi}'_1 \mid \boldsymbol{\psi}'_0) p(\boldsymbol{\xi}_1 \mid \boldsymbol{\xi}_0) \end{array} \right) d\boldsymbol{\Xi}_k && (5.8) \end{aligned}$$

Due to the recursive dependence of $\boldsymbol{\xi}_k$ on $\boldsymbol{\xi}_{k-1}$, it is not possible to write equation (5.8) in a time recursive form as in equation (5.1). A further approximation is thus made, that the nuisance parameters, $\boldsymbol{\xi}_k$, are independent across frames, so that $p(\boldsymbol{\xi}_k \mid \boldsymbol{\xi}_{k-1}) = p(\boldsymbol{\xi}_k)$, where $p(\boldsymbol{\xi}_k)$ represents some prior distribution. Using this independence assumption across frames, it now becomes possible to split the integral into two separate integrals, over $d\boldsymbol{\Xi}_{k-1}$ and $d\boldsymbol{\xi}_k$, restoring the sequential update nature of the equation as shown below:

$$\begin{aligned} p(\boldsymbol{\Psi}'_k \mid \mathbf{X}_k) &\propto \int p(\boldsymbol{\Psi}_{k-1} \mid \mathbf{X}_{k-1}) d\boldsymbol{\Xi}_{k-1} \times p(\boldsymbol{\psi}'_k \mid \boldsymbol{\psi}'_{k-1}) \int p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) p(\boldsymbol{\xi}_k) d\boldsymbol{\xi}_k && (5.9) \\ &\propto p(\boldsymbol{\Psi}'_{k-1} \mid \mathbf{X}_{k-1}) \times p(\boldsymbol{\psi}'_k \mid \boldsymbol{\psi}'_{k-1}) p(\mathbf{x}_k \mid \boldsymbol{\psi}'_k) && (5.10) \end{aligned}$$

where, as illustrated earlier in equation (4.33) (on page 88):

$$p(\mathbf{x}_k \mid \boldsymbol{\psi}'_k) = \int p(\mathbf{x}_k \mid \boldsymbol{\psi}_k) p(\boldsymbol{\xi}_k) d\boldsymbol{\xi}_k \tag{5.11}$$

In the particle filter context, the marginalised posterior distribution $p(\boldsymbol{\Psi}'_k \mid \mathbf{X}_k)$ can be approx-

imated using a set of N weighted samples as:

$$p(\mathbf{\Psi}'_k \mid \mathbf{X}_k) \approx \sum_{i=1}^{N} {w'_k}^{(i)} \delta(\mathbf{\Psi}'_k - {\mathbf{\Psi}'_k}^{(i)}) \tag{5.12}$$

where the *marginalised weight* ${w'_k}^{(i)}$ is given by:

$$w'_k = \frac{p({\mathbf{\Psi}'_k}^{(i)} \mid \mathbf{X}_k)}{q({\mathbf{\Psi}'_k}^{(i)} \mid \mathbf{X}_k)} \tag{5.13}$$

When the integral in equation (5.11) can be performed analytically, a marginalised form of the particle filter is obtained leading to what is referred to as the *MPF* in this thesis. In the case of the general linear model (GLM) described in Chapter 4, it has been shown in Section 4.4.4 that marginalisation of the amplitude parameters and noise variance is possible, giving rise to the marginalised distribution $p(\mathbf{x}_k \mid \boldsymbol{\psi}'_k)$ as given by equation (4.40).

Further simplification of the weight update equation is possible when the transition prior is used as the sampling distribution, *i.e.*, $q(\boldsymbol{\psi}'_k|\boldsymbol{\psi}'_{k-1}) = p(\boldsymbol{\psi}'_k|\boldsymbol{\psi}'_{k-1})$, yielding:

$$w'_k \propto w'_{k-1} p(\mathbf{x}_k \mid \boldsymbol{\psi}'_k) \tag{5.14}$$

The following section discusses application of the MPF to the problem of sequential frequency estimation. The algorithm uses a RJMCMC move (introduced in Section 3.3.2) to detect the unknown number of tones and harmonics. The construction of the specific moves will be discussed here. Sections 5.4 and 5.5 will subsequently compare the MPF and the RBPF within the sequential frequency estimation context.

5.2 Frequency Estimation using the Marginalised Particle Filter

The MPF described in Section 5.1 can be used to perform sequential frequency estimation allowing marginalisation of certain parameters. Following the notation used in Section 4.4, the state vectors are defined as $\boldsymbol{\psi}'_k = \{P_k, \mathbf{S}_k, \mathbf{H}_k, \delta_k^2\}$, and $\boldsymbol{\xi}_k = \{\mathbf{a}_k, \sigma^2_{(e),k}\}$, where P_k denotes the number of tones, $\mathbf{S}_k$ denotes the parameters of the polynomial basis, $\mathbf{H}_k$ is the vector denoting the number of harmonics of each tone in the state and δ_k^2 is an indication of the SNR [29]. $\boldsymbol{\xi}_k$ denotes the parameters which are to be marginalised, consisting of the vector of amplitudes

$\mathbf{a}_k$ and the observation noise variance $\sigma^2_{(e),k}$. While these parameters are not necessarily independent across frames, this assumption is crucial to the marginalisation. The impact of this assumption will be examined later in this chapter. The priors, hyper-parameters and likelihood function associated with the model have been previously described in Section 4.4.2.

The following section will discuss the transition priors and state update equations before examining the RJMCMC framework for estimating the number of tones and harmonics.

5.2.1 State Update Equations

As described in Chapter 4, the update for the polynomial basis parameters is given by:

$$\mathbf{s}_k^{(p)} = \Delta_s \mathbf{s}_{k-1}^{(p)} + \mathbf{v}_{(s),k}^{(p)}, \qquad \forall p \in \{1, \ldots, P_{k-1}\} \tag{5.15}$$

where Δ_s is defined in equation 4.19. The definition of Δ_s requires that the time between updates Δt is specified. While equation (4.32) describes the initial prior distribution for the parameter δ_k^2, the state update for δ_k^2 is given by:

$$\log(\delta_k^2) = \log(\delta_{k-1}^2) + v_{(\delta),k} \tag{5.16}$$

where $v_{(\delta),k} \sim \mathcal{N}(0, \sigma^2_{(\delta)})$. This form of the update is necessary to ensure that δ_k^2 is always positive. Since the amplitude and noise variance will be marginalised out, they do not require to be specified in the state update equations.

The remaining parameters, *i.e.*, the number of tones P_k, and the number of harmonics $\mathbf{H}_k$ are updated using a RJMCMC move rather than a JMS framework (as in Dubois and Davy's RBPF algorithm described in Section 3.5) due to the advantages associated with the RJMCMC framework. The following section will discuss the particle filtering algorithm, and Section 5.3 will cover application of the reversible jump sampler to the estimation of the number of tones and harmonics.

5.2.2 Particle Filtering Algorithm

Implementation of the algorithm using the MPF depends on the availability of the marginalised likelihood function, $p(\mathbf{x}_k \mid \boldsymbol{\psi}'_k)$, and assumes that the marginalised parameters $\boldsymbol{\xi}_k$ are independent across frames. The marginalised likelihood function can be analytically evaluated for the

GLM as discussed in Section 4.4.4. Consequently, implementation of the MPF for sequential frequency tracking is straightforward and the resulting filter is almost identical to the SIR filter with the exception that the marginalised likelihood function is used to calculate the weights.

The state update equations illustrate how the polynomial basis parameters are updated. The number of tones and harmonics are updated using separate reversible jump moves. The construction and discussion of these moves is covered in Section 5.3, and the steps involved in the implementation of the MPF are listed in Algorithm 4.

Algorithm 4 Tracking a signal with varying, unknown number of tones and harmonics using the MPF and the two jump method.

Initialisation ($k = 0$):
$(P_0^{(i)}, \mathbf{S}_0^{(i)}, \delta_0^{2(i)}) \sim p(P, \mathbf{S}, \delta^2) \mid_{i=1}^{N}$ where i is the particle index, and N is the number of particles used.

for $k = 1 : k_{max}$ **do**
  Process-update:
  $\mathbf{S}_k^{(i)} \sim p(\mathbf{S} \mid P_{k-1}^{(i)}, \mathbf{S}_{k-1}^{(i)}) \mid_{i=1}^{N}$
  $\mathbf{M}_k^{(i)} \sim p(\mathbf{M}_k \mid P_{k-1}^{(i)}, \mathbf{M}_{k-1}^{(i)}) \mid_{i=1}^{N}$
  $\delta_k^{2(i)} \sim p(\delta^2 \mid (\delta_{k-1}^2)^{(i)}) \mid_{i=1}^{N}$
  Evaluate importance weight using equations (4.40) and (5.14):
  $\hat{w}_k'^{(i)} = {w'_{k-1}}^{(i)} p(\mathbf{x}_k \mid {\boldsymbol{\psi}'_k}^{(i)})$
  Normalise weights:
  ${w'_k}^{(i)} = T^{-1} \hat{w}_k'^{(i)} \mid_{i=1}^{N}$ where $T = \sum_{i=1}^{N} \hat{w}_k'^{(i)}$
  Obtain parameter estimates (equations (5.17)-(5.19)).
  Resample:
  $[[{\boldsymbol{\psi}'_k}^{(i)} w_k^{(i)}] = \text{RESAMPLE}[\{{\boldsymbol{\psi}'_k}^{(i)}, w_k^{(i)}\}_{i=1}^{N}]$
  Determine number of tones using the reversible jump move:
  ${\boldsymbol{\psi}'_k}^{(i)} = \text{REVERSIBLE JUMP}[\{{\boldsymbol{\psi}'_k}^{(i)}\}_{i=1}^{N}]$ (page 101)
  Determine the number of harmonics for each tone
  for $i = 1 : N$ **do**
    ${\boldsymbol{\psi}'_k}^{(i)} = \text{HARMONIC RJ}[\{{\boldsymbol{\psi}'_k}^{(i)}, p\}_{p=1}^{P_k^{(i)}}]$ (page 105
  end for
end for

5.2.3 Obtaining Parameter Estimates from the Filter

The particle filter produces an approximation of the multidimensional distribution over the signal parameters for each frame. Each particle represents an unordered set of frequencies (or polynomials). As a result, taking a sample mean does not necessarily yield a meaningful estimate of the instantaneous frequency content of the signal. As explained in [70], a marginal

maximum *a posteriori* estimate can be computed which corresponds to the frequency components "most represented" in the particles. Such an estimate requires sorting of the frequencies so that the marginal MMSE estimate of the parameters is meaningful [70].

In this work, a *max weight* estimator [104] has been used which assumes that the particle with the maximum weight is representative of the parameters at that instant. This estimator is extremely simple, but is suitable for multi-modal distributions and, at the same time, does not require sorting of the parameters as is necessary for the marginal MMSE estimate. The estimate of the parameters, $\hat{\boldsymbol{\psi}}_k'$, can be written as:

$$\hat{i}_k = \arg_i \max w_k^{(i)} \tag{5.17}$$

$$\hat{\boldsymbol{\psi}}_k' = \boldsymbol{\psi}_k'^{(i)} \tag{5.18}$$

For the MPF, since the amplitudes have been marginalised out, it is possible to obtain an estimate of the amplitudes $\hat{\mathbf{a}}_k$using the maximum likelihood estimate given the parameters $\hat{\boldsymbol{\psi}}_k'$ as:

$$\hat{\mathbf{a}}_k = (\hat{\mathbf{G}}_k^T \hat{\mathbf{G}}_k)^{-1} \hat{\mathbf{G}}_k^T \mathbf{x}_k \tag{5.19}$$

where $\hat{\mathbf{G}}_k$ is constructed as a function of the $\hat{\boldsymbol{\psi}}_k'$ parameters.

5.3 Estimation of Tones and Harmonics using Reversible Jump Methods

Reversible jump samplers [30] are useful in cases where the number of parameters that need to be determined is one of the parameters itself. Such a case arises here where the number of tones and harmonics, as well as their frequencies needs to be determined.

Andrieu and Doucet [29], Copsey *et al*. [33] and Davy *et al*. [38] discuss this estimation problem in a batch offline scenario. In this section, the process as described by Davy *et al*. [38] will be described which will be referred to here as the two-jump move (TJM) to indicate the use of two sets of reversible jump moves.

The first reversible jump is applied to determine the number of tones present while the second, applied to each tone in turn, is used to determine the number of associated harmonics. The use of multiple reversible jump moves can add undesirable computational burden to the algorithm.

Consequently, Chapter 6 will introduce and compare two alternative reversible jump schemes which attempt to reduce the complexity of the algorithm by making certain simplifying assumptions.

5.3.1 Estimation of the Number of Tones

The reversible jump sampler is applied independently to each particle at time k. The number of components tracked by each particle is then updated according to a birth, death, or update move. The birth and death moves respectively increase and decrease the number of components, while an update move leaves the number of components unaltered. The probability of selecting a particular move is defined by the birth, death and update probabilities, $\{b_t, d_t, u_t\}$ respectively [29, 35]:

$$\begin{aligned} b_t &= c_t \cdot \min\left\{\frac{p(P_k+1)}{p(P_k)}, 1\right\} \\ d_t &= c_t \cdot \min\left\{\frac{p(P_k-1)}{p(P_k)}, 1\right\} \\ u_t &= 1 - b_t - d_t \end{aligned} \tag{5.20}$$

where c_t is a constant (as in [35], the value of c_t is set to 0.5) and the subscript t is used to indicate move probabilities associated with the number of tones. The probabilities for the number of tones are determined according to a Poisson distribution, $p(P_k) = p(P_k \mid \lambda_t)$, as described in Section 4.4.2.

Setting c_t to a large value (0.5 here) ensures that new states will be proposed for a large number of particles in the filter, thus facilitating the search over the sample space. This is in contrast with the JMS setup described previously in Section 3.5 where the majority of particles undergo an update-only move. The reversible jump sampler does not suffer from the use of a large value of c_t since the new states may be rejected if they are associated with a low acceptance probability. Section 6.2 will examine the birth and death probabilities for the TJM and compare it with alternative sampling schemes developed in the next chapter.

Once a particular move has been selected, the acceptance ratio for that move needs to be calculated. Let $\boldsymbol{\psi}_k$ denote the current state, and let the superscript $\star$ be used to denote the new proposed state. The posterior distribution is obtained as the product of the priors listed in Section 4.4.2 and the marginalised likelihood function given by equation (4.40).

For a birth move, $P_k^\star = P_k + 1$, a new polynomial needs to be appended to the existing

parameter vector. The new polynomial is sampled from a distribution so that $\mathbf{s}^+ \sim g(\mathbf{s})$, and the proposed vector is denoted as $\mathbf{S}_k^\star = [\mathbf{S}_k; \mathbf{s}^+]$. The proposal distributions for the polynomial basis is chosen as a multidimensional uniform distribution over suitable limits as stated in Section 4.4.2. Additionally, the overtone deviation associated with the fundamental frequency of the new tone is restricted to zero as discussed in Section 4.4.2. The proposal distribution for the birth move, $d(\boldsymbol{\psi}_k^\star \mid \boldsymbol{\psi}_k)$, can then be written as [35]:

$$d(\boldsymbol{\psi}_k^\star \mid \boldsymbol{\psi}_k) = p(P_k + 1 \mid \lambda_t) \times g(\mathbf{s}^+) \tag{5.21}$$

$$\propto \frac{\lambda_t^{P_k+1}}{(P_k+1)!} \times g(\mathbf{s}^+) \tag{5.22}$$

Similarly, for the inverse death move, from P_k+1 to P_k, a polynomial $(\mathbf{s}^-)$ is randomly chosen for elimination from the set of existing components yielding the proposal distribution:

$$d(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_k^\star) \propto \frac{\lambda_t^{P_k}}{P_k!} \times \frac{1}{P_k+1}$$

Before evaluating the acceptance ratio, the Jacobian $\mathbf{J}$ needs to be calculated for the different moves. The Jacobian arises due to the *dimension-matching requirement* which necessitates a bijection between $\boldsymbol{\psi}_k$ and $\boldsymbol{\psi}_k^\star$ [30]. Green [30] describes in detail how the moves may be constructed while maintaining this dimension-matching requirement. A simple example will be considered here to illustrate the process and show how the Jacobian is calculated. This can then be extended to the birth and death moves for determining the number of tones and harmonics.

Consider a move from a state with a single frequency component to a state with two frequency components using a birth move. Let $(P = 1, f)$ denote the current state, and $(P = 2, f^{(1)}, f^{(2)})$ denote the proposed new state. The dimension-matching criterion involves generating a continuous random variable u independently of f. The move is then performed between the augmented state (f, u) and $(f^{(1)}, f^{(2)})$. In the case of a birth move, the transition can be written as:

$$f \rightarrow f^{(1)}$$

$$u \rightarrow f^{(2)}$$

such that the new state is formed as a deterministic function of f and u. The Jacobian is then

evaluated as $\mathbf{J} = \left| \dfrac{\partial(f^{(1)}, f^{(2)})}{\partial(f, u)} \right|$:

$$\mathbf{J} = \begin{vmatrix} \dfrac{\partial f^{(1)}}{\partial f} & \dfrac{\partial f^{(1)}}{\partial u} \\ \dfrac{\partial f^{(2)}}{\partial f} & \dfrac{\partial f^{(2)}}{\partial u} \end{vmatrix} = 1 \tag{5.23}$$

When constructing more complicated moves such as split or merge, the Jacobian may not evaluate to 1, however, for the birth and death moves considered in this and the following sections, the Jacobian evaluates to unity and will be ignored.

Using equation (5.23), the acceptance ratios for the moves are then evaluated as:

$$\begin{aligned} r_{birth} &= \frac{p(\boldsymbol{\psi}_k^\star \mid \mathbf{x}_k)}{p(\boldsymbol{\psi}_k \mid \mathbf{x}_k) g(\mathbf{s}^+) \lambda_t} \\ r_{death} &= r_{birth}^{-1} \end{aligned} \tag{5.24}$$

where, as stated in Section 4.4.4, the prior distribution $g(\mathbf{s}^+)$ is chosen as a uniform distribution, and marginalised likelihood is given by equation (4.40).

The distribution $g(\cdot)$ can be tailored to the application in question. In the simplest case, the distribution can be specified as a uniform prior over suitable limits, for example, the frequency can lie between zero and the Nyquist frequency. For the specific case when the sampling distribution $g(\mathbf{s}^+)$ is identical to the uniform prior on the parameters, the birth move ratio simplifies to:

$$r_{birth} = \frac{p(\mathbf{x}_k \mid \boldsymbol{\psi}_k^\star)}{p(\mathbf{x}_k \mid \boldsymbol{\psi}_k)} \frac{1}{P_k + 1} \tag{5.25}$$

As an alternative to the uniform prior, it is possible to use the Fourier transform as a sampling distribution for new frequencies. Dubois and Davy [70] discuss other algorithms for constructing efficient importance densities. Under such circumstances, the acceptance ratio needs to be evaluated using equation (5.24) since it does not simplify as shown above.

By applying the jump move to each particle after importance sampling, the filter is able to track the time-varying number of tones. Algorithm 5 lists the steps involved in the reversible jump move. Once the number of tones has been updated, another reversible jump move is applied to *each* tone to determine the number of harmonics of that tone. This is done through a similar birth/death, or multiply/divide move [38] which will be discussed next.

Algorithm 5 Steps involved in the reversible jump sampler for determining an unknown number of tones.

procedure REVERSIBLEJUMP($\boldsymbol{\psi}$)
  With probability:
    b_t, select a harmonic birth move.
    d_t, select a harmonic death move.
    $u_t = 1 - b_h - d_h$, select an update only move.
  if birth move **then**
    Propose
      $P_k^\star = P_k + 1$
      $\mathbf{S}_k^\star = [\mathbf{S}_k; \mathbf{s}^+]$ where $\mathbf{s}^+ \sim g(\mathbf{s}^+)$
      $M^{(p)\star} = 0$
  else if death move **then**
    Propose
      $P_k^\star = P_k - 1$
      $\mathbf{S}_k^\star = \mathbf{S}_k \setminus \mathbf{s}^-$
      $\mathbf{M}^\star = \mathbf{M} \setminus M^-$
  end if
  Evaluate acceptance probability α.
  Sample $u \sim \mathcal{U}_{[0,1]}$
  if $u \leq \alpha$ **then**
    accept new state $\boldsymbol{\psi}^\star = [P_k^\star, \mathbf{S}_k^\star, \mathbf{M}_k^\star$.
  else
    retain old state $\boldsymbol{\psi}$.
  end if
end procedure

5.3.2 Harmonic Birth and Death Moves

Birth and death moves allow the algorithm to track a time-varying number of tones in the signal. Through a similar process, the number of harmonics of the p^{th} tone is updated using an n-increase/decrease move [38]. For an n-increase move, n is sampled from a discrete uniform distribution such that [38]:

$$n \sim \mathcal{U}_d[0, \max(H_{max} - H^{(p)}, H^{(p)} - H_{min})] \tag{5.26}$$

$$H_k^{(p)\star} = H_k^{(p)} + n, \quad H_k^{(p)\star} \in [H_{min}, H_{max}] \tag{5.27}$$

where $\mathcal{U}_d[a, b]$ denotes a discrete uniform distribution over the limit $[a, b]$. The birth move then adds n additional harmonics to the tone.

A Poisson prior has been specified for the number of harmonics and the probabilities of selecting birth and death moves for the harmonics update are given by [38]:

$$\begin{aligned} b_h &= \gamma c_h \cdot \min\left\{\frac{p(H^{(p)} + n \mid \lambda_h)}{p(H^{(p)} \mid \lambda_h)}, 1\right\} \\ d_h &= \gamma c_h \cdot \min\left\{\frac{p(H^{(p)} \mid \lambda_h)}{p(H^{(p)} + n \mid \lambda_h)}, 1\right\} \end{aligned} \tag{5.28}$$

where c_h is a constant and the subscript h is used to denote moves associated with determination of the number of harmonics. Similar to the previously described reversible jump move for the tones, the value of c_h is set to 0.5. The constant γ controls the probability of selecting either the birth/death moves or the multiply/divide moves, which are described in the next section.

Addition of new harmonics may require sampling of the overtone deviation parameters if they are included in the model. The conditional proposal distributions are defined as:

$$d(\boldsymbol{\psi}_k^\star \mid \boldsymbol{\psi}_k) = p(H^{(p)} + n \mid \lambda_h) g(\mathbf{M}^+) \tag{5.29}$$

$$d(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_k^\star) = p(H^{(p)} \mid \lambda_h) \tag{5.30}$$

where $\mathbf{M}^+ = [\mu_1^+, \ldots, \mu_n^+]^T$, and $g(\mathbf{M}^+)$ denotes the sampling distribution for the additional overtone deviation parameters. The distribution $g(\mathbf{M}^+)$ is assumed to be identical to the prior distribution for simplicity. The acceptance ratio for the n-increase move then evaluates to:

$$h_{birth} = \frac{p(\mathbf{x} \mid \boldsymbol{\psi}^\star)}{p(\mathbf{x} \mid \boldsymbol{\psi})} \tag{5.31}$$

with the converse death move having an acceptance ratio as $h_{death} = h_{birth}^{-1}$.

5.3.3 Harmonic Multiply and Divide Moves

The harmonic multiply and divide moves [38] are necessary to resolve ambiguities involving the fundamental tone. For a signal consisting of a single tone with multiple harmonics, the algorithm is required to determine which frequency corresponds to the fundamental tone and which frequencies are harmonics. In the absence of multiply and divide moves, it is possible for the algorithm to get caught in local maxima where either a sub-harmonic or a higher harmonic is tracked rather than the true fundamental tone. The multiply and divide moves serve to resolve such ambiguities. An example of such a local maxima trap will be illustrated in Section 6.3.

A multiply move works by multiplying the the fundamental frequency of a tone by two, thus halving the number of harmonics in the signal. Conversely, the divide move halves the frequency of the fundamental tone and doubles the number of harmonics. Figures 5.1 and 5.2 illustrate the effect of the multiply and divide moves.

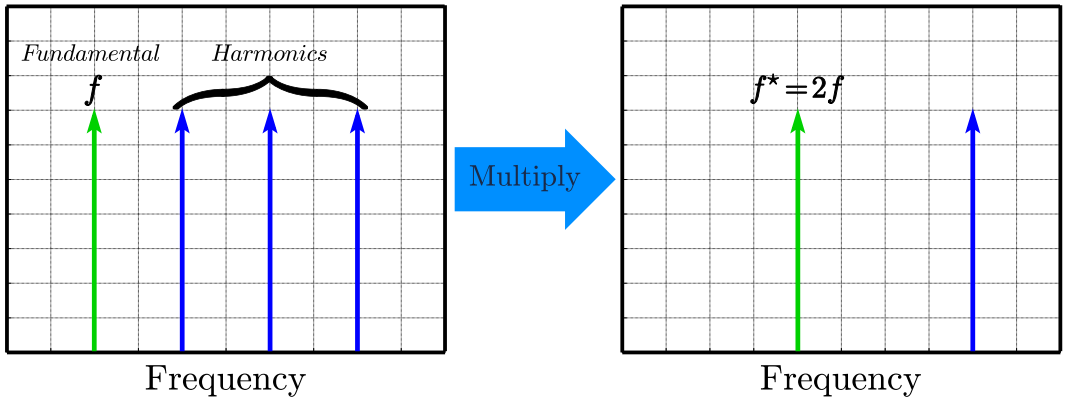


Figure 5.1: *Effect of a multiply move on a component.*

The probability of selecting a multiply or divide move is respectively given by [38]:

$$m_h = \frac{1-\gamma}{2} \tag{5.32}$$

$$v_h = \frac{1-\gamma}{2} \tag{5.33}$$

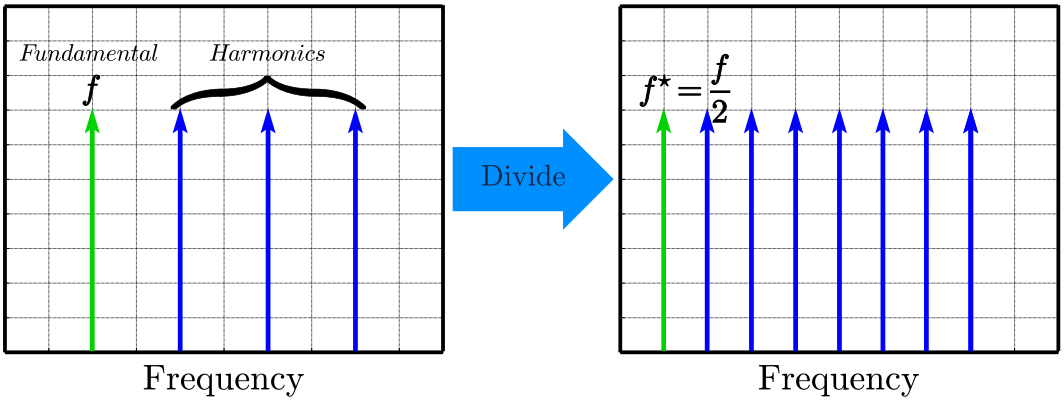


Figure 5.2: *Effect of a divide move on a component.*

For a divide move, the conditional proposal distributions are given by:

$$d(\boldsymbol{\psi}_k^\star \mid \boldsymbol{\psi}_k) = p(H^{(p)^\star} \mid \lambda_h) g(\mathbf{M}^+) \tag{5.34}$$

$$d(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_k^\star) = p(H^{(p)} \mid \lambda_h) \tag{5.35}$$

where $g(\mathbf{M}^+)$ denotes the sampling distribution for the overtone deviation parameters introduced by the additional harmonics. Since $g(\mathbf{M}^+)$ is assumed identical to the prior distribution, the acceptance ratio for the divide move is given by:

$$h_{divide} = \frac{p(\mathbf{x} \mid \boldsymbol{\psi}^\star)}{p(\mathbf{x} \mid \boldsymbol{\psi})} \tag{5.36}$$

The multiply move is the inverse move of the divide move and the acceptance ratio is obtained as $h_{multiply} = h_{divide}^{-1}$.

Using one reversible jump move to detect the number of tones and another for each tone to estimate the number of harmonics makes this approach computationally demanding. This motivates the search for less complex approaches to perform this task. Two such approaches will be considered in the next chapter. Algorithm 6 lists the steps involved in the harmonic reversible jump move for detecting an unknown number of harmonics for the p^{th} tone.

Algorithm 6 Steps involved in the harmonic reversible jump move for the p^{th} tone

procedure HARMONIC RJ($\boldsymbol{\psi}, p$)
  Sample $n \sim \mathcal{U}_d[0, \max(H_{max} - H^{(p)}, H^{(p)} - H_{min})]$
  With probability:
    b_h, select a harmonic birth move.
    d_h, select a harmonic death move.
    m_h, select a harmonic multiply move.
    v_h, select a harmonic divide move.
    $u_h = \gamma(1 - b_h - d_h)$, select an update only move.
  if birth move **then**
    Propose
      $H^{(p)\star} = H^{(p)} + n$
      $M^{(p)\star} = [M^{(p)}; M^{+}]$
  else if death move **then**
    Propose
      $H^{(p)\star} = H^{(p)} - n$
      $M^{(p)\star} = [\mu_1^{(p)}, \ldots, \mu_{H^{(p)}-n}^{(p)}]^T$
  else if multiply move **then**
    Propose
      $f^{(p)\star} = 2f^{(p)}$
      $H^{(p)\star} = \text{floor}\left(\frac{H^{(p)}}{2}\right)$
      $M^{(p)\star} = [\mu^{(p,2)}, \mu^{(p,4)}, \ldots, \mu^{(p,H^{(p)\star})}]^T$
  else if divide move **then**
    Propose
      $f^{(p)\star} = \dfrac{f^{(p)}}{2}$
      $H^{(p)\star} = 2H^{(p)}$
      $M^{+} = [\mu^{(+,1)}, \mu^{(+,2)}, \ldots, \mu^{(+,H^{(p)})}]^T$
      $M^{(p)\star} = [\mu^{(+,1)}, \mu^{(p,1)}, \mu^{(+,2)}, \mu^{(p,2)}, \ldots, \mu^{(+,H^{(p)})}, \mu^{(p,H^{(p)})}]^T$
  end if
  Evaluate acceptance probability α.
  Sample $u \sim \mathcal{U}_{[0,1]}$
  if $u \leq \alpha$ **then**
    accept new state $\boldsymbol{\psi}^{\star}$.
  else
    retain old state $\boldsymbol{\psi}$.
  end if
end procedure

5.4 Computational Complexity of the RBPF and MPF

The RBPF and the MPF use slightly different methods to achieve marginalisation of parameters. This section compares the complexity of both marginalisation techniques. While this comparison is placed within the sequential frequency estimation context, the comparison contrasts the complexity of the RBPF and MPF constructs themselves, rather than the corresponding frequency estimation algorithms described in Sections 3.5 and 5.2. The complexity arising from the use of the RJMCMC moves is thus ignored here.

The time-complexity of both the RBPF and MPF can be judged by examining the matrix operations which are required in the Kalman filter and the corresponding likelihood functions used in each method. To obtain the complexity of both algorithms, the complexity of matrix multiplication, matrix inversion and calculation of the matrix determinant need to be known. For the sake of simplicity, the complexity is considered when a naïve approach is adopted in evaluation of each of these matrix operations, or in other words, no form of optimisation is considered for the operations.

For the case of matrix multiplication, multiplication of two rectangular matrices is considered. Multiplying a $p \times q$ matrix with a $q \times r$ matrix results in a $p \times r$ matrix and the associated complexity is $\mathcal{O}(pqr)$ [105]. For a $n \times n$ square matrix, calculation of a matrix inverse can be performed in $\mathcal{O}(n^3)$ time [105, 106]. Similarly, evaluation of the determinant of a $n \times n$ square matrix has a time complexity of $\mathcal{O}(n^3)$ [106].

In the frequency estimation context considered here, the complexity of the algorithms is determined by the window length L, and the number of tones and harmonics given by O_k. In evaluating the complexity, it is assumed that the length of the window is much larger than the number of tones and harmonics, *i.e.*, $L \gg O_k$.

The complexity involved in the Kalman filter is illustrated in Table 5.1. From the table, the dominant contribution to the complexity arises from evaluation of the inverse of the covariance matrix $\mathbf{S}_k$, and the complexity of the Kalman filter is then given by $\mathcal{O}(L^3)$.

It is also necessary to evaluate the complexity of the likelihood function used in the RBPF. The

Kalman filter equation	Computational complexity (dominating term only)
$\boldsymbol{\mu}_{k\|k-1} = \mathbf{A}\boldsymbol{\mu}_{k-1}$	$\mathcal{O}(O_k^2)$
$\boldsymbol{\Sigma}_{k\|k-1} = \mathbf{A}\boldsymbol{\Sigma}_{k-1}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T$	$\mathcal{O}(O_k^3)$
$\mathbf{S}_k = \mathbf{C}\boldsymbol{\Sigma}_{k\|k-1}\mathbf{C}^T + \mathbf{D}\mathbf{D}^T$	$\mathcal{O}(L^3)$
$\hat{\mathbf{x}}_{k\|k-1} = \mathbf{C}\boldsymbol{\mu}_{k\|k-1}$	$\mathcal{O}(O_k L)$
$\mathbf{R}_k = \mathbf{S}_k^{-1}$	$\mathcal{O}(L^3)$
$\boldsymbol{\mu}_k = \boldsymbol{\mu}_{k\|k-1} + \boldsymbol{\Sigma}_{k\|k-1}\mathbf{C}^T\mathbf{R}_k(\mathbf{x}_k - \hat{\mathbf{x}}_{k\|k-1})$	$\mathcal{O}(L^2)$
$\boldsymbol{\Sigma}_k = \boldsymbol{\Sigma}_{k\|k-1} - \boldsymbol{\Sigma}_{k\|k-1}\mathbf{C}^T\mathbf{R}_k\mathbf{C}\boldsymbol{\Sigma}_{k\|k-1}$	$\mathcal{O}(O_k L^2)$

Table 5.1: *Computational complexity involved in the Kalman filter equations.*

likelihood function used in the RBPF is given by [70]:

$$p(\mathbf{x}_k|\boldsymbol{\Psi}'_k, \mathbf{X}_{k-1}) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{S}_k) \tag{5.37}$$

$$= \frac{1}{(2\pi)^{\frac{L}{2}}\sqrt{|\mathbf{S}_k|}} \exp\left\{-\frac{1}{2}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T\mathbf{S}_k^{-1}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\right\} \tag{5.38}$$

The most expensive operation present in the above likelihood function arises from the calculation of the matrix determinant, which has a complexity of $\mathcal{O}(L^3)$. The matrix inverse needs to be evaluated for the Kalman filter equations and does not need to be recalculated here. After the determinant, the next most expensive operation comes from matrix multiplication with a complexity of $\mathcal{O}(L^2)$. The complexity of the RBPF can thus be regarded as $\mathcal{O}(L^3)$.

In the case of the MPF, the main source of complexity arises from the likelihood function since there is no additional Kalman filter step involved. The likelihood function used in the MPF is derived from marginalisation of parameters in the GLM (see Section 4.4) and the corresponding likelihood function is reproduced here:

$$p(\mathbf{x}_k \mid \boldsymbol{\psi}'_k) \propto \frac{\left(\mathbf{x}_k^T\mathbf{x}_k + 2\alpha_e - \mathbf{x}_k^T\mathbf{G}_k\mathbf{F}_k^{-1}\mathbf{G}_k^T\mathbf{x}_k\right)^{-\left(\frac{L}{2}+\beta_e\right)}}{\delta_k^{2O_k}\sqrt{|\mathbf{F}_k|}} \tag{5.39}$$

where $\mathbf{F} = (\mathbf{G}_k^T\mathbf{G}_k + \delta_k^{-2}\mathbb{I}_{2O_k})$.

While the marginalised likelihood given above also possesses a matrix inverse and determinant, the complexity of both operations is $\mathcal{O}(O_k^3)$ since $\mathbf{F}_k$ is a $2O_k \times 2O_k$ matrix. The dominant term in the complexity of the marginalised likelihood function is seen to come from evaluation

of the $\mathbf{F}_k$ matrix and has a complexity of $\mathcal{O}(O_k^2 L)$ due to the term $\mathbf{G}_k^T \mathbf{G}_k$, where $\mathbf{G}_k$ is a $L \times 2O_k$ matrix. The complexity of the MPF is thus given by the dominant term of $\mathcal{O}(O_k^2 L)$.

When the RBPF and the MPF are now compared, it is seen that the RBPF possesses a much higher complexity of $\mathcal{O}(L^3)$ versus $\mathcal{O}(O_k^2 L)$ for the MPF.

When a more generic context is considered rather than the specific frequency estimation context here, the length of the observation vector $\mathbf{x}_k$ will be given by L and the number of columns in the $\mathbf{G}_k$ matrix will reflect the number of basis functions used in the model.

It is thus seen that the MPF is able to show computational advantages over the RBPF. A comparison between the performance of the RBPF and the MPF will be carried out in the next section within the context of sequential frequency estimation where it will be shown that the performance of an MPF based algorithm (see Section 5.2) is nearly identical to the RBPF based method developed by Dubois and Davy [70] (see Section 3.5).

5.5 Comparison between RBPF and MPF Frequency Estimators

Chapter 3 described a sequential frequency estimation algorithm developed by Dubois and Davy which is based on the RBPF [70]. An alternative, new marginalised particle filter and frequency estimation algorithm was developed in this chapter. This section provides a comparison between the two algorithms to show their strengths and weaknesses.

5.5.1 Overview of Experimental Setup

Neglecting any differences between the basis used in the signal models of the two methods, *i.e.*, sinusoidal vs. polynomial chirps, the major difference between the methods is in the construction of the marginalisation process and the means by which the number of tones and harmonics is detected. This section looks at the difference in performance between the two methods, first examining the effect of the different marginalisation techniques, and then reviewing the difference between the JMS and RJMCMC frameworks.

The experiments may be separated into three sets. The first set aims to measure the effect of the different marginalisation techniques. This is performed by restricting the estimation to a single tone. The second set compares the ability of the algorithms to accurately detect the

correct number of tones when a single sinusoid is present for varying window lengths. The experiments conclude with the final set which compares the performance of the algorithms on a multicomponent signal.

In all the comparisons, both algorithms are run with the same number of particles. In the frequency estimation technique based on the RBPF, the observation noise variance is assumed fixed and known. As discussed in Section 4.4.2, the hyper-parameters for the MPF are selected so as to produce uninformative priors. Additionally, in both the signal models, the number of harmonics is restricted to one, *i.e.*, the tones do not possess any higher harmonics.

In examining the effect of the marginalisation techniques, differences in performance of the two algorithms may be expected to arise due to the independence assumption used in deriving the MPF, *i.e.*, the nuisance parameters are assumed independent across frames. The aim of these experiments is to quantify the differences in the level of performance of the two algorithms.

5.5.2 Constrained Monocomponent Estimation

It is first established that both algorithms produce a similar frequency mean squared error (MSE) under constrained estimation (the number of estimated components is restricted to one). The squared frequency error at the k^{th} instant is measured as $e_k^f = (f_k - \hat{f}_k)^2$ where f_k is the true IF and $\hat{f}_k$ is the IF estimate. The test signal used is a single sinusoid in noise, with a normalised constant frequency of 0.23. Figure 5.3 shows the frequency MSE when the same signal is examined under differing SNR conditions using 100 particles and averaged over 100 runs. The figure shows that both algorithms offer near identical performance under these constrained conditions.

Table 5.2 summarises the frequency MSE of the two different methods. Although the figure indicates that the performance is similar, the table shows that the RBPF outperforms the MPF. This is conditional, however, on accurate knowledge of the observation noise variance. The table shows that underestimating this noise variance causes the performance of the RBPF to drop. In the case of the MPF, the noise variance parameter has been marginalised out of the distribution and thus a similar degradation in performance is not observed.

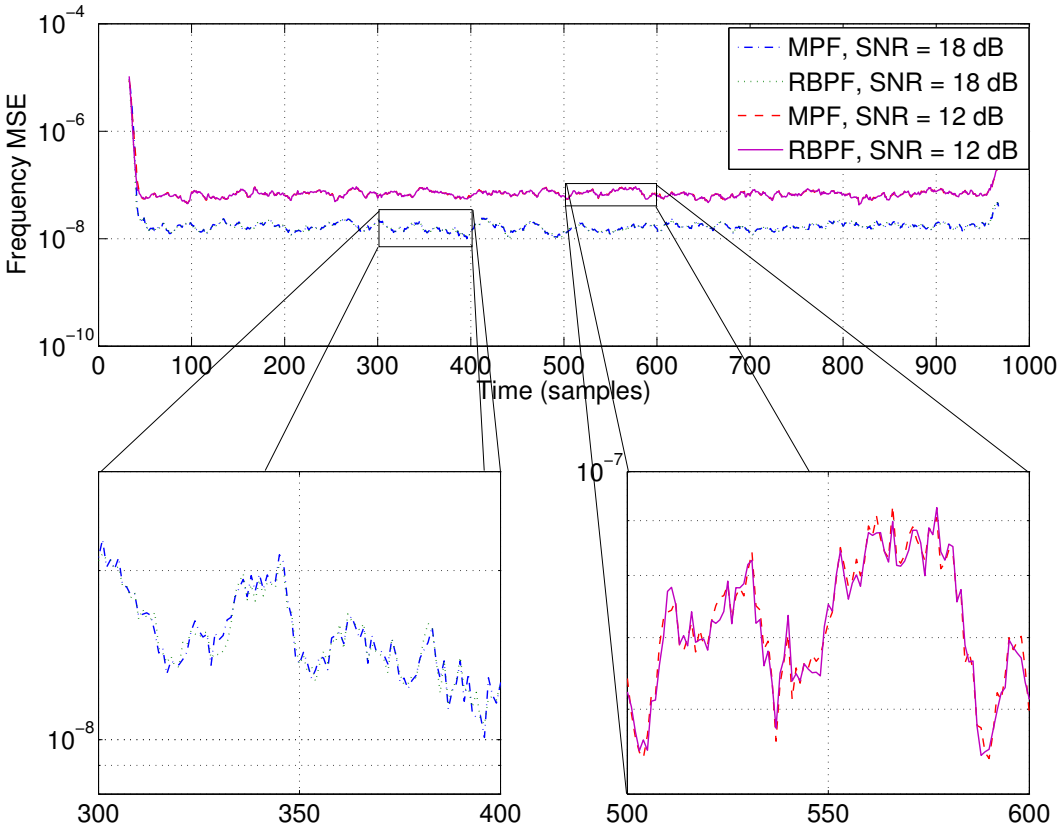


Figure 5.3: *The figure shows a comparison of the frequency MSE for the two marginalisation methods. It is clear that under constrained settings, both methods offer similar levels of performance.*

SNR	Frequency MSE for RBPF (10^{-7})	Frequency MSE for MPF (10^{-7})
18	0.3720	0.4561
12	0.9338	0.9597
09	1.5386	1.5972
09*	1.7708	

Table 5.2: *Comparison of the frequency MSE for different SNR. The SNR marked * indicates the use of noise variance underestimated by a factor of* 2 *in the filter.*

5.5.3 Unconstrained Monocomponent Estimation

These simulations are carried out again under unconstrained conditions, *i.e.*, both the number of tones and their frequencies are estimated. As a measure of estimation error, the *model order* MSE is used since it has been shown that both algorithms provide similar performance under constrained detection. The model order is defined as the total number of components present in the signal. The model order squared error is obtained as $e_k^O = (O_k - \hat{O}_k)^2$, where $O_k = \sum_{p=1}^{P_k} H_k^{(p)}$ denotes the true model order and $\hat{O}_k$ is the estimated model order.

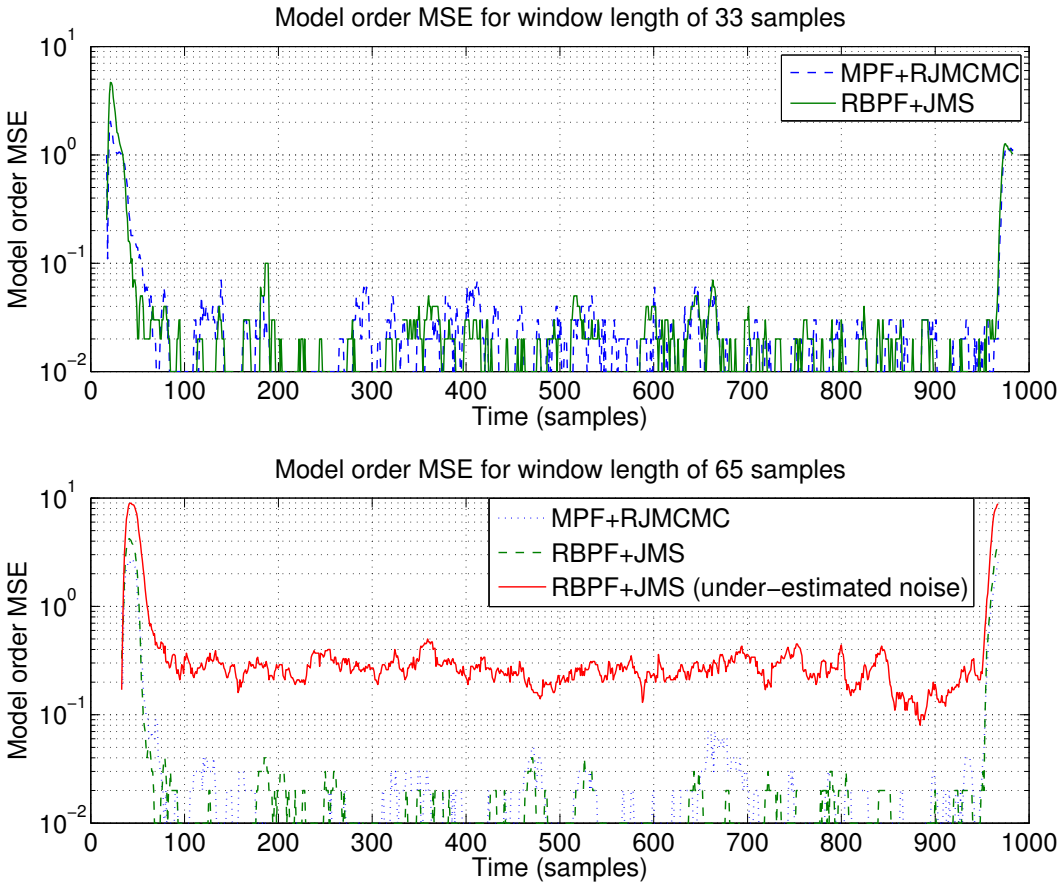


Figure 5.4: *The model order MSE of both methods is similar, however, the RBPF is sensitive to the observation noise.*

Figure 5.4 shows the resulting model order MSE for both methods for window lengths of 33 and 65 samples. "RBPF+JMS" refers to the algorithm developed by Dubois and Davy while "MPF+RJMCMC" refers to the algorithm developed here. The large error at the start and end of the estimates is due to the presence of some degree of amplitude modulation which is present only at the start and end of the signal. Table 5.3 lists the mean of the model order MSE for the

Algorithm	Window Length	Model Order MSE
MPF+RJMCMC	33	0.01749
RBPF+JMS	33	0.01569
MPF+RJMCMC	65	0.01384
RBPF+JMS	65	0.00988
RBPF+JMS (under-estimated noise)	65	0.26526
MPF+RJMCMC	129	0.00758
RBPF+JMS	129	0.00455

Table 5.3: *Comparison of model order MSE for the algorithms with varying window length.*

methods, as evaluated between time samples 100 and 900. This ensures that error arising due to the amplitude modulation is ignored.

From the table, it is seen that the method developed by Dubois and Davy performs better, consistently producing a lower model order MSE. However, this improved performance is dependent on knowledge of the observation noise variance (assumed known), and the quality of the estimates is sensitive to this value. Figure 5.4 shows an example of the model order MSE when the observation noise is assumed to be half the true value, indicated as "RBPF+JMS (under-estimated noise)". Under this erroneous assumption, the model order MSE is seen to increase significantly.

5.5.4 Multicomponent Estimation

The final set of comparisons involves the use of a multicomponent signal. While Figure 5.4 shows the results from tracking a single sinusoid in noise, Figures 5.5-5.8 present the results from tracking a signal with multiple sinusoids with components starting and ending at different times. Figure 5.5 shows the time domain and time-frequency representation of the test signal used (the spectrogram of the signal is generated using a window length of 128 samples). The signal has an approximate SNR of 30 dB. Similar to the previous set of results, the model order MSE is used as a measure of algorithm performance.

Figure 5.6 compares the model order MSE for the two algorithms. The error is averaged over 100 runs using 100 particles for each run. A window length of 33 samples is used in these

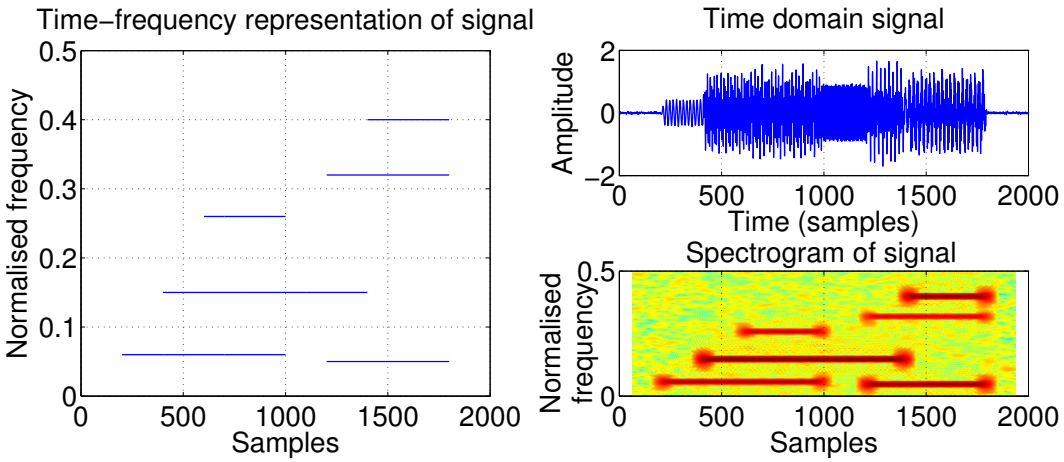


Figure 5.5: *A multicomponent signal is used to measure the effect of the marginalisation techniques as well as compare the JMS and reversible jump sampler in detecting new frequency components in the signal.*

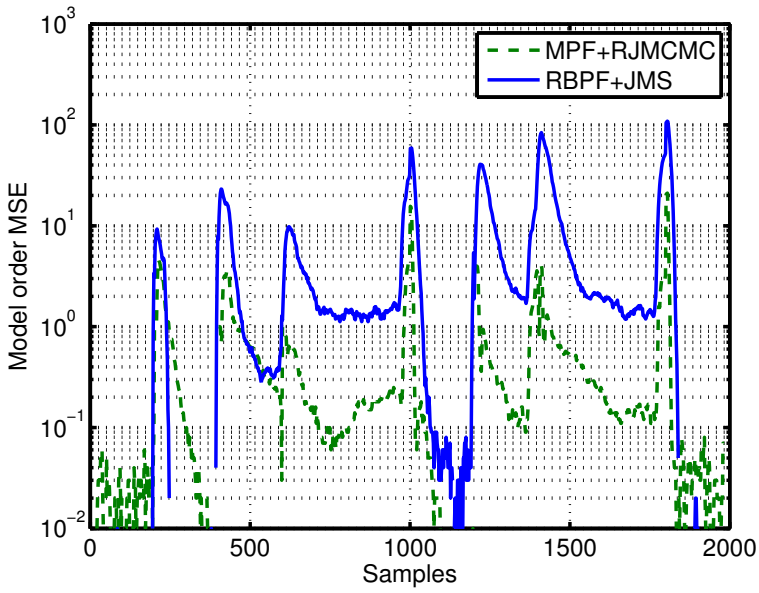


Figure 5.6: *Comparison of the model order MSE for both algorithms.*

simulations. The figure shows that despite the observation noise being unknown, the MPF is able to perform well. In comparing the results, it is seen that the algorithm developed by Dubois and Davy does not suffer from spurious detections at the start or end of the signal when no components are present. However, in other sections of the signal, when multiple components are present, the MPF-based algorithm is able to determine the number of signal subcomponents with greater precision due to the use of the RJMCMC sampler. The error is plotted on a semilog scale, and hence the error is not plotted when the average error is zero. Examples of the estimation results are shown in Figures 5.7 and 5.8.

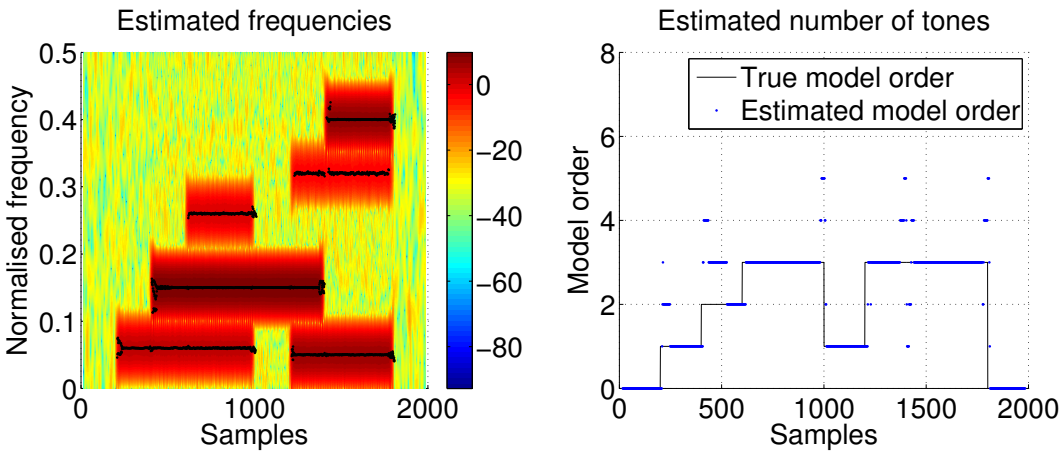


Figure 5.7: *Frequency and model order estimates using analytical marginalisation with a reversible jump move for model order estimation.*

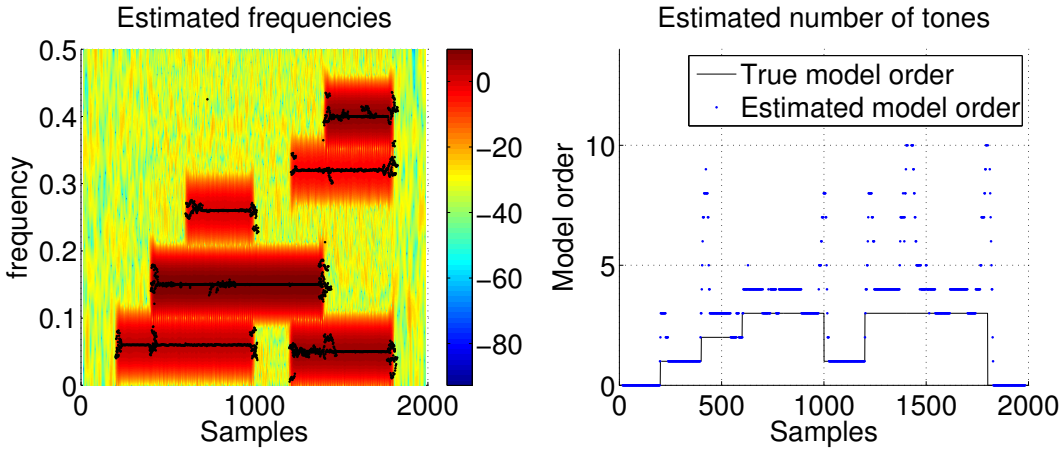


Figure 5.8: *Frequency and model order estimates using Rao-Blackwellisation with a JMS for model order estimation.*

5.5.5 Conclusions

The set of experiments carried out here have been used to compare the RBPF and the MPF algorithms for sequential frequency estimation. The constrained monocomponent test serves to compare only the RBPF and MPF constructs by eliminating the need to estimate a varying number of tones through either a JMS or RJMCMC framework.

The results show that the MPF performs only slightly worse than the RBPF when used to track the frequency of a monocomponent signal using the same number of particles. These results are encouraging, since the MPF possesses a significantly lower complexity than the equivalent RBPF algorithm. Additionally, the MPF algorithm assumes that the observation noise is not known, in contrast with the RBPF algorithm considered here.

Having established that there is very little difference between the performance of the MPF and the RBPF, the multicomponent signal is used to measure the difference between the use of the JMS and the RJMCMC sampler for estimation of the number of tones in a signal. The results show that the "RBPF+JMS" algorithm performs well when there is either only noise, or a single frequency component. However, In the presence of multiple components, the "MPF+RJMCMC" algorithm shows a much lower model order error. This can be attributed to the reversible jump sampler which ensures that good particles are not lost as a result of transitions to low-likelihood states, thus maintaining a lower error bound.

5.6 Summary

This chapter developed an alternative marginalisation scheme for particle filters. The MPF developed here offers an alternative to the well established RBPF.

Section 5.2 developed an algorithm for sequential frequency estimation. This algorithm, in contrast with the one developed by Dubois and Davy [70] (described in Section 3.5), relies on a reversible jump framework to determine the time-varying number of tones and harmonics. The disadvantage of using a JMS with predefined transition probabilities is that it is possible for particles to transition to states with low probability, and to prevent this, low transition probabilities may be associated with birth and death transitions. Consequently, this limits the extent to which the posterior distribution is explored.

Although more computationally expensive, RJMCMC methods offer a better method for deter-

mining the number of tones and harmonics. To perform a transition to a different dimension new state, the new state is first proposed and is accepted or rejected according to the acceptance probability. As discussed in [90], the use of this accept/reject mechanism prevents good particles from being lost to states associated with low likelihoods, while simultaneously allowing the algorithm to fully explore the state space by allowing high transition probabilities. Section 5.2 examines the algorithm developed by Davy *et al.* in [38] for estimation of the number of tones and harmonics, showing how this can be used in the particle filtering framework.

Section 5.4 compared the complexity of the MPF and the RBPF when applied to the linear and Gaussian class of problems. It was shown that under these circumstances, the RBPF has a significantly higher computational complexity of $\mathcal{O}(L^3)$ while the MPF has a complexity of $\mathcal{O}(O_k^2 L)$ where L is the length of the observation vector and the number of basis functions used is given by $2O_k$.

Section 5.5 compares the use of a JMS and a reversible jump sampler for determining the time-varying number of tones in a signal. The results reflect the advantages of the reversible jump method in the reduced MSE (of the number of detected tones) associated with this method.

The chapter concludes with a comparison between the sequential frequency estimation algorithm based on the RBPF as described in Section 3.5, and the MPF as described in Section 5.2. The comparison shows that the MPF performs favourably against the RBPF despite the independence assumption used in the marginalisation process. The RBPF is capable of better performance, however, this is dependent on knowledge of the observation noise. The MPF is able to circumvent this requirement by marginalisation of the observation noise variance and using uninformative priors to indicate a complete lack of knowledge.

With regards to estimation of the number of tones and harmonics, the reversible jump sampler is seen to provide better performance than a simple JMS. Despite the benefits, the reversible jump sampler can prove to be computationally expensive, especially when first applied to each particle to determine the number of tones, and then subsequently to each tone of each particle to determine the number of harmonics. The following chapter aims to address this issue of high computational complexity and examines two alternatives to the two-jump move (TJM) described here to reduce the complexity without impacting the estimation performance.

Chapter 6
Estimation of Harmonics in Signals

Introduction

Chapter 4 established a signal model for dealing with a multi-tone signal containing multiple harmonics. Chapter 5 subsequently derived a new sequential frequency estimation algorithm based on the marginalised particle filter (MPF) which uses a reversible jump sampler to estimate the time-varying number of harmonics. The reversible jump sampler used there, referred to as the two-jump move (TJM) in this work, is based on the batch offline algorithm developed by Davy *et al*. [38]. The TJM works by applying one reversible jump move to determine the number of tones present in a signal, followed by a reversible jump move applied to each tone to determine the number of associated harmonics.

The application of multiple reversible jump moves in the TJM can be computationally expensive and the aim of this chapter is to develop alternative reversible jump schemes. The use of a limited number of samples in the particle filter limits the ability of the algorithm to explore the entire state space and, consequently, the design of more *efficient* reversible jump moves becomes important. Alternatives are thus considered which attempt to reduce the computational complexity of the algorithm without sacrificing the performance of the estimator.

In the following section, two alternatives to the TJM will be introduced based on the assumption of a *slowly varying signal*. In the context here, slowly varying implies that only a single tone may undergo any change across subsequent time instants. Such an assumption may not hold for all signals. For example, a complex piece of music may involve multiple notes from different instruments changing simultaneously. In contrast with this, naturally occurring signals and in particular, bat echolocation calls, are not expected to exhibit such sudden changes to the signal structure over short periods of time, thus satisfying the assumption of a slowly varying signal.

In developing the algorithms in this chapter, it is noted that these are not the only methods for sampling the multidimensional state space. The methods described here are two specific cases considered in this work. Alternative moves can be proposed to detect the time-varying number

of tones and harmonics in different ways and may be more appropriate for different types of signals.

In addition to the reversible jump schemes developed here, Section 6.4 at the end of this chapter, establishes the estimation accuracy of the inharmonicity parameter within the sequential context. These limits may restrict the usefulness of the method in detecting inharmonicity in signals and it is essential that these be examined before application to bat echolocation calls as considered in the next chapter.

6.1 Reversible Jump Sampling Schemes: Alternatives to the Two-Jump Move

The TJM developed by Davy *et al*. [38] uses a reversible jump framework to detect an unknown number of tones and harmonics in the batch offline scenario. Chapter 5 described how this algorithm can be adapted to perform the same task in a sequential particle filter framework.

The TJM is computationally expensive due to the presence of two sets of reversible jump moves. For a state with P_k number of tones, $P_k + 1$ reversible jump moves are carried out – once to establish the number of tones, and then once for each tone in the state. To perform the move, each reversible jump move requires the evaluation of the acceptance probability, and consequently, the posterior density ratio. It is clear that for a large number of tones, the computational complexity becomes exorbitant. This section will develop two reversible jump schemes based on the previously mentioned assumption of a slowly varying signal in an attempt to reduce the complexity associated with the multiple reversible jump moves.

Before discussing the particulars of the different sampling schemes, some details of the reversible jump sampler are reproduced here for convenience. The reversible jump move is applied independently to each particle in the k^{th} frame. The number of components tracked by each particle is then updated according to a birth, death or update move in the case of tones, and birth, death, multiply, divide or update move for harmonics. The probabilities of these moves is determined according to the prior distributions for the number of tones and harmonics described in Section 4.4.2. Once a particular move has been selected, a new set of parameters

$\boldsymbol{\psi}^{\star}$ is proposed and the acceptance ratio is evaluated according to the expression [30]:

$$r = \frac{p(\boldsymbol{\psi}_k^{\star} \mid \mathbf{x}_k)}{p(\boldsymbol{\psi}_k \mid \mathbf{x}_k)} \frac{d(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_k^{\star})}{d(\boldsymbol{\psi}_k^{\star} \mid \boldsymbol{\psi}_k)} \mathbf{J} \tag{6.1}$$

where $d(\cdot \mid \cdot)$ denotes the conditional proposal distribution for the parameters. The new set of parameters is then accepted according to the acceptance probability $\alpha = \min(1, r)$. As discussed in Section 5.3, the Jacobian $\mathbf{J}$ evaluates to 1 for the moves considered in this thesis. The move probabilities and distributions will be discussed for the specific schemes in the following sections.

6.1.1 The Modified Two-Jump Method (MTJM)

The modified two-jump move (MTJM) may be considered a variation of the TJM described in the previous chapter. By incorporating the assumption of a slowly varying signal, *i.e.*, only a single tone may undergo any change to the number of harmonics at any instant, the MTJM is able to reduce the complexity otherwise associated with the TJM.

In the MTJM, the first reversible jump move remains unchanged and is used to detect the number of tones in the signal. The second reversible jump move, however, is applied to a single tone selected at random, to determine the number of associated harmonics.

Once the p^{th} tone has been selected, as in the case of the TJM, the selected tone may undergo a harmonic birth/death or multiply/divide move. In the case of the birth or death moves, the number of harmonics $H_k^{(p)}$ is increased or decreased respectively using an n-increase or decrease move, where n is sampled according to equations (5.26) and (5.27), reproduced here:

$$n \sim \mathcal{U}_d[0, \max(H_{max} - H^{(p)}, H^{(p)} - H_{min})] \tag{6.2}$$

$$H_k^{(p)\star} = H_k^{(p)} + n, \quad H_k^{(p)\star} \in [H_{min}, H_{max}] \tag{6.3}$$

The conditional proposal distributions are written as:

$$d(\boldsymbol{\psi}_k^{\star} \mid \boldsymbol{\psi}_k) = \frac{1}{P_k} p(H_k^{(p)\star} \mid \lambda_h) g(\mathbf{M}^{+}) \tag{6.4}$$

$$d(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_k^{\star}) = \frac{1}{P_k} p(H_k^{(p)} \mid \lambda_h) \tag{6.5}$$

where $g(\mathbf{M}^{+})$ is the sampling distribution for the vector of inharmonicity parameters. For the

case of the multiply or divide moves, the number of harmonics associated with the new state gets halved or doubled respectively.

Similar to the TJM, the prior distribution is chosen for the sampling distribution $g(\mathbf{M}^+)$ resulting in the acceptance ratio:

$$h_{birth} = h_{divide} = \frac{p(\mathbf{x} \mid \boldsymbol{\psi}_k^\star)}{p(\mathbf{x} \mid \boldsymbol{\psi}_k)} \tag{6.6}$$

The acceptance ratio for the death and multiply moves are consequently obtained as $h_{death} = h_{birth}^{-1}$ and $h_{multiply} = h_{divide}^{-1}$.

The limitations of the MTJM are most noticeable when considering a signal with multiple tones. If the number of harmonics of all the tones change simultaneously, then this model will be unable to track the change instantaneously. Instead, there will be a lag associated with tracking the harmonics of some of the tones. Using a large number of particles will improve performance but cannot completely eliminate this shortcoming. The implementation of the MTJM within the particle filter framework is described in Algorithm 7.

6.1.2 The Combined Jump Method (CJM)

The combined jump move (CJM) attempts to combine the two separate reversible jump moves into a single move to further reduce the complexity by eliminating the second move of the MTJM altogether. The move is constructed by combining elements of the tone and harmonic reversible jump moves.

A birth move consists of adding either a new tone with n harmonics, or adding n harmonics to an existing tone. A specific case will also be considered later, where during addition of a new tone, n is constrained to one, *i.e.*, the new tone will not have higher harmonics. This allows the algorithm to add either a new tone (with no additional harmonics), or additional harmonics to an existing tone. A death move involves the removal of n harmonics from an existing tone, so that a tone is removed if it has exactly n harmonics. The multiply and divide moves are identical to those described previously, resulting in half and twice the number of harmonics respectively.

The probability of selecting a birth or death move is slightly different from the probabilities used in the previous methods. Prior to evaluating the move probabilities, the following random

Algorithm 7 Tracking a signal with an unknown, varying number of tones and harmonics using the modified two-jump method

Initialisation ($k = 0$):
$(P_0^{(i)}, \mathbf{f}_0^{(i)}, \mathbf{M}_0^{(i)}, \delta_0^{2(i)}) \sim p(P, \mathbf{f}, \mathbf{M}, \delta^2) \mid_{i=1}^{N}$ where i is the particle index, and N is the number of particles used.

1: **for** $k = 1 : k_{max}$ **do**
2: Process-update:

$$\begin{aligned} \mathbf{f}_k^{(i)} &\sim p(\mathbf{f} \mid P_{k-1}^{(i)}, \mathbf{f}_{k-1}^{(i)}, \mathbf{c}_{k-1}^{(i)}) \mid_{i=1}^{N} \\ \mathbf{M}_k^{(i)} &\sim p(\mathbf{M}_k \mid P_{k-1}^{(i)}, \mathbf{M}_{k-1}^{(i)}) \mid_{i=1}^{N} \\ \delta_k^{2(i)} &\sim p(\delta^2 \mid (\delta_{k-1}^2)^{(i)}) \mid_{i=1}^{N} \end{aligned}$$

3: Evaluate importance weight:
$\hat{w}_k^{(i)} \propto w_{k-1}^{(i)} p(\mathbf{x}_k \mid \boldsymbol{\psi}_k^{(i)})$
4: Normalise weights:
$w_k^{(i)} = T^{-1} \hat{w}_k^{(i)} \mid_{i=1}^{N}$ where $T = \sum_{i=1}^{N} \hat{w}_k^{(i)}$
5: Resample:
$[\boldsymbol{\psi}_k^{(i)} w_k^{(i)}] = \text{RESAMPLE}[\{\boldsymbol{\psi}_k^{(i)}, w_k^{(i)}\}_{i=1}^{N}]$
6: Determine number of tones using the reversible jump move:
$\boldsymbol{\psi}_k^{(i)} = \text{REVERSIBLE JUMP}[\{\boldsymbol{\psi}_k^{(i)}\}_{i=1}^{N}]$ (See Algorithm 5 on page 101)
7: Determine the number of harmonics for a single tone selected at random
8: **for** $i = 1 : N$ **do**
$p \sim \mathcal{U}_d[1, P_k^{(i)}]$
Sample $\boldsymbol{\psi}_k^{(i)} = \text{HARMONIC RJ}[\{\boldsymbol{\psi}_k^{(i)}, p\}]$ (See Algorithm 6 on page 105)
9: **end for**
10: **end for**

variables are drawn:

$$\begin{aligned} P_b &\sim \mathcal{U}_d[1, P_k + 1] \\ n_b &\sim \mathcal{U}_d[1, H_{max} - H_k^{(P_b)}] \end{aligned} \tag{6.7}$$

where P_b is the index of the tone selected to be updated, n_b is the number of harmonics to be added to the tone and P_k is the number of tones in the current state. When $P_b = P_k + 1$, this indicates addition of a new tone, and the number of associated harmonics is defined as $H_k^{(P_b)} = 0$. In the case of a birth move, when a new tone is added, $P_k^\star = P_k + 1$ and $H_k^{(P_k+1)} = 0$, otherwise, $P_k^\star = P_k$. Similar to P_b and n_b which are associated with birth moves, P_d and n_d are drawn for evaluating the death move:

$$\begin{aligned} P_d &\sim \mathcal{U}_d[1, P_k] \\ n_d &\sim \mathcal{U}_d[1, H_k^{(P_d)}] \end{aligned} \tag{6.8}$$

where P_d is the index of the tone selected for the n-decrease move and n_d is the number of harmonics to be removed. When $n_d = H_k^{(P_d)}$, this indicates complete removal of that tone and proposes a move from $P_k + 1$ to $P_k^\star = P_k$ number of tones, otherwise, $P_k^\star = P_k$.

The move probabilities are then evaluated using the prior distributions for the number of tones and harmonics as:

$$\begin{aligned} b_c &= \gamma c_c \cdot \min\left\{ \frac{p(P_k^\star \mid \lambda_t) p(H_k^{(P_b)} + n_b \mid \lambda_h)}{p(P_k \mid \lambda_t) p(H_k^{(P_b)} \mid \lambda_h)}, 1 \right\} \\ d_c &= \gamma c_c \cdot \min\left\{ \frac{p(P_k^\star \mid \lambda_t) p(H_k^{(P_d)} - n_d \mid \lambda_h)}{p(P_k \mid \lambda_t) p(H_k^{(P_d)} \mid \lambda_h)}, 1 \right\} \\ m_c &= \frac{1-\gamma}{2} \\ v_c &= \frac{1-\gamma}{2} \end{aligned} \tag{6.9}$$

where the constant $c_c = 0.5$ and γ determines the probability of selecting the birth/death or the multiply/divide moves.

The conditional proposal distributions are written as:

$$d(\boldsymbol{\psi}_k^\star \mid \boldsymbol{\psi}_k) = p(P_k^\star \mid \lambda_t) p(H_k^{(p)\star} \mid \lambda_h) \frac{1}{P_k + 1} g(\mathbf{s}^+) g(\mathbf{M}^+) \tag{6.10}$$

$$d(\boldsymbol{\psi}_k \mid \boldsymbol{\psi}_k^\star) = p(P_k \mid \lambda_t) p(H_k^{(p)} \mid \lambda_h) \frac{1}{P_k^\star} \tag{6.11}$$

where $g(\mathbf{s}^+)$ and $g(\mathbf{M}^+)$ are sampling distributions for the added parameters and are the same as the prior distributions. While the number of tones may change in the birth and death moves, there is no change for multiply and divide moves, $P_k^\star = P_k$. The acceptance ratio then evaluates to:

$$h_{birth} = h_{divide} = \frac{p(\mathbf{x} \mid \boldsymbol{\psi}_k^\star)(P_k + 1)}{p(\mathbf{x} \mid \boldsymbol{\psi}_k)P_k^\star} \tag{6.12}$$

Using a combined move designed in this manner may be expected to be slower at tracking changes in a multi-tone signal when there are simultaneous changes to both the number of tones and harmonics. However, assuming that these variations occur slowly on the scale of the window length, such an assumption may be deemed acceptable and combining the two jumps serves to further reduce the computational complexity of the jump moves.

Algorithm 8 Steps involved in the combined jump move for estimation of the number of tones and harmonics

procedure COMBINED RJ($\boldsymbol{\psi}$)
    Sample P_b, n_b, P_d, n_d according to equations 6.7 and 6.8
    With probability:
        b_c, select a birth move.
        d_c, select a death move.
        m_c, select a harmonic multiply move.
        $v_{(c)}$, select a harmonic divide move.
        $u_c = \gamma(1 - b_c - d_c)$, select an update only move.
     to obtain new state parameters $\boldsymbol{\psi}^\star$.
    Evaluate acceptance probability α.
    Sample $u \sim \mathcal{U}_{[0,1]}$
    if $u \leq \alpha$ **then**
        accept new state $\boldsymbol{\psi}^\star$.
    else
        retain old state $\boldsymbol{\psi}$.
    end if
end procedure

6.2 Examination of Birth and Death Probabilities

Before comparing the performance of the various sampling schemes, the probability of selecting birth and death moves is examined. These move probabilities will shed some light on the performance of the algorithms since the construction of the moves influences how the sample space of tones and harmonics is explored.

Algorithm 9 Tracking a signal with an unknown, varying number of tones and harmonics using the combined jump method

Initialisation ($k = 0$):
$(P_0^{(i)}, \mathbf{f}_0^{(i)}, \mathbf{M}_0^{(i)}, \delta_0^{2(i)}) \sim p(P, \mathbf{f}, \mathbf{M}, \delta^2) \mid_{i=1}^{N}$ where i is the particle index, and N is the number of particles used.

1: **for** $k = 1 : k_{max}$ **do**

2: Process-update:

$$\begin{aligned} \mathbf{f}_k^{(i)} &\sim p(\mathbf{f} \mid P_{k-1}^{(i)}, \mathbf{f}_{k-1}^{(i)}, \mathbf{c}_{k-1}^{(i)}) \mid_{i=1}^{N} \\ \mathbf{M}_k^{(i)} &\sim p(\mathbf{M}_k \mid P_{k-1}^{(i)}, \mathbf{M}_{k-1}^{(i)}) \mid_{i=1}^{N} \\ \delta_k^{2(i)} &\sim p(\delta^2 \mid (\delta_{k-1}^2)^{(i)}) \mid_{i=1}^{N} \end{aligned}$$

3: Evaluate importance weight:
$\hat{w}_k^{(i)} \propto w_{k-1}^{(i)} p(\mathbf{x}_k \mid \boldsymbol{\psi}_k^{(i)})$

4: Normalise weights:
$w_k^{(i)} = T^{-1} \hat{w}_k^{(i)} \mid_{i=1}^{N}; \;\; T = \sum_{i=1}^{N} \hat{w}_k^{(i)}$

5: Resample:
$\boldsymbol{\psi}_k^{(i)} = \text{RESAMPLE}[\{\boldsymbol{\psi}_k'^{(i)}, w_k^{(i)}\}_{i=1}^{N}]$

6: Determine number of tones and harmonics using the reversible jump move:
$\boldsymbol{\psi}_k^{(i)} = \text{COMBINED RJ}[\{\boldsymbol{\psi}_k^{(i)}, \}_{i=1}^{N}]$

7: **end for**

The TJM and MTJM may be seen as treating the number of tones and harmonics as being independent and the addition of harmonics is unrelated to the addition of tones. The CJM attempts to merge these two steps and consequently associates these steps.

The difference between the proposal probabilities is demonstrated using an example considering only birth and death moves for both tones and harmonics. The hyper-parameters for the distributions on the number of tones and harmonics are set as $\lambda_t = 1$ and $\lambda_h = 5$. The limits for the number of tones and harmonics are specified as $P_{min} = 0$, $P_{max} = 15$, $H_{min} = 1$ and $H_{max} = 10$.

Considering only birth and death moves for tones, the probability of selecting a birth, death or update move in the case of the TJM and MTJM is given by equation (5.20), reproduced here:

$$\begin{aligned} b_t &= c_t \cdot \min\left\{\frac{p(P_k + 1)}{p(P_k)}, 1\right\} \\ d_t &= c_t \cdot \min\left\{\frac{p(P_k - 1)}{p(P_k)}, 1\right\} \\ u_t &= 1 - b_t - d_t \end{aligned} \tag{6.13}$$

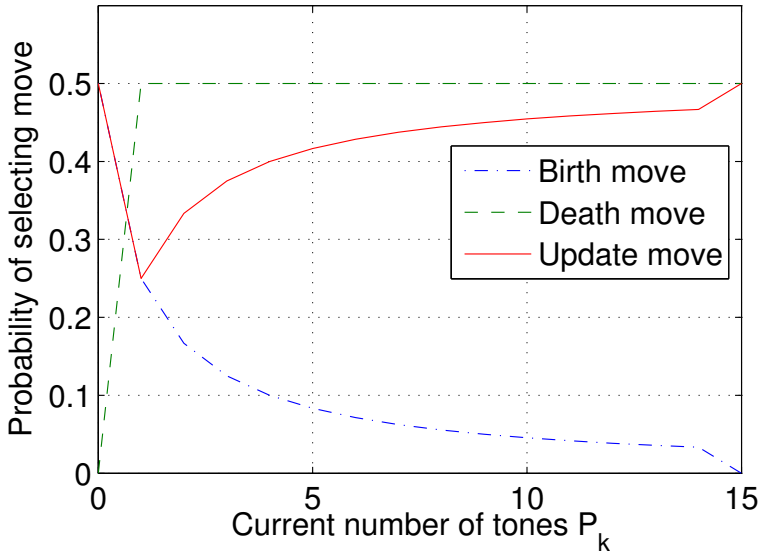


Figure 6.1: *Probability of selecting a birth, death or update move for the tone reversible jump move.*

where the number of tones in the current state is given by P_k.

Using these expressions, Figure 6.1 shows the probability of selecting a tone birth, death and update move for the TJM and MTJM. From the figure, it can be seen that as the number of tones in the state increases, the probability of selecting a birth move decreases while that of the death move approaches a constant value. The constant c_t dictates the proportion of particles that undergo a birth or death move, with the remainder subjected to an update-only move.

For the case of a harmonic birth/death move, the probability of proposing an n-increase or decrease move is given by equation (5.28), reproduced here:

$$\begin{aligned} b_h &= \gamma c_h \cdot \min\left\{\frac{p(H^{(p)}+n \mid \lambda_h)}{p(H^{(p)} \mid \lambda_h)}, 1\right\} \\ d_h &= \gamma c_h \cdot \min\left\{\frac{p(H^{(p)} \mid \lambda_h)}{p(H^{(p)}+n \mid \lambda_h)}, 1\right\} \end{aligned} \tag{6.14}$$

with $\gamma = 1$ here, so that multiply and divide moves are ignored.

These harmonic birth and death probabilities are evaluated on a grid and shown in Figures 6.2 and 6.3. The figures show the probability of selecting a harmonic birth and death move going from H_k to $H_k^\star$ harmonics. The figures suggest a region of preference which is governed by

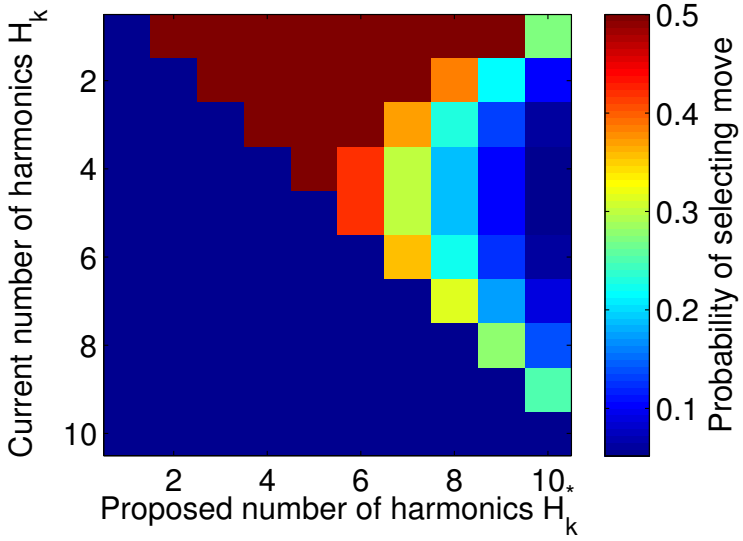


Figure 6.2: *Probability of going from H_k to $H_k^\star$ harmonics for a selected tone using a harmonic birth move.*

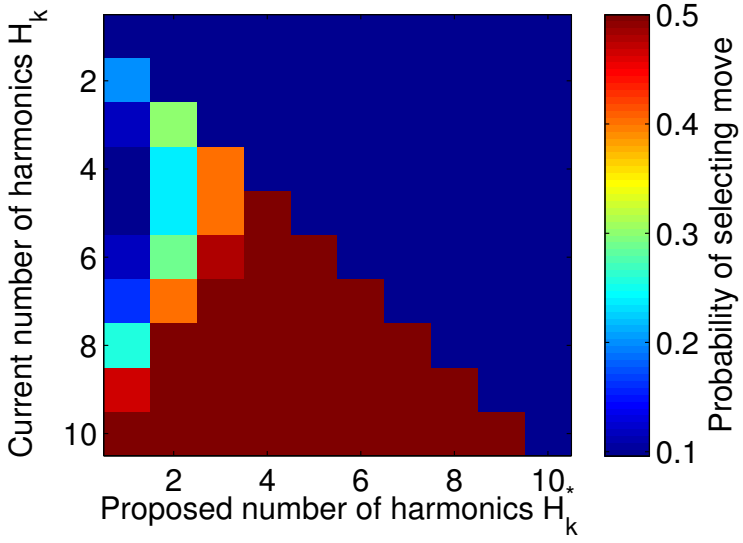


Figure 6.3: *Probability of going from H_k to $H_k^\star$ harmonics for a selected tone using a harmonic death move.*

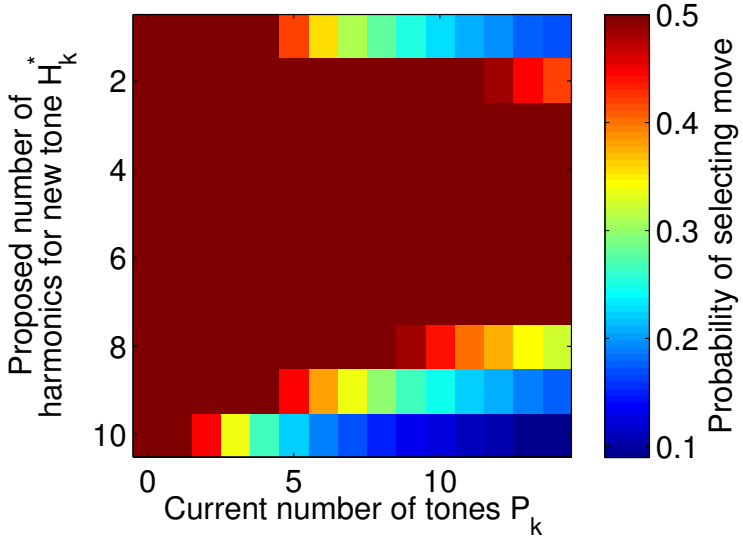


Figure 6.4: *Probability of adding a new tone with $H_k^\star$ harmonics using the CJM when there are P_k tones present.*

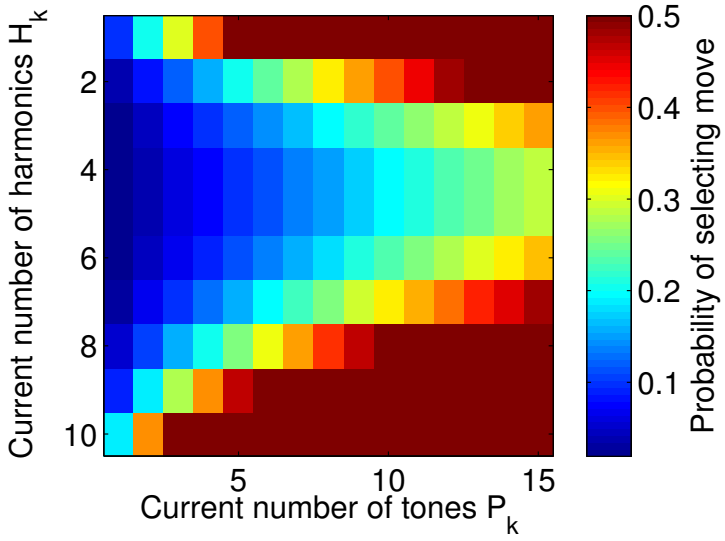


Figure 6.5: *Probability of deleting a tone with $H_k^\star$ harmonics using the CJM when there are P_k tones present.*

the hyper-parameter λ_h. Birth and death moves for the number of harmonics will be favoured if they help the state approach this region. A similar trend is seen in Figure 6.1 where there is a tendency to move towards a particular number of tones governed by the value of λ_t.

The case for the CJM is slightly different. This difference arises since the probabilities for selecting a particular move are calculated *after* a tone has been selected. Setting $\gamma = 1$ in equation (6.9), the birth and death probabilities are evaluated on a grid and shown in Figures 6.4 and 6.5.

Figure 6.4 shows the probability of adding a new tone with $H_k^\star$ harmonics when the current number of tones is P_k. Figure 6.5 shows the probability of deleting a tone (and all its harmonics) when there are P_k tones and the selected tone has H_k harmonics. For the example used here, the figures show that in the presence of a large number of tones, selecting a tone with either very many or very few harmonics results in an increased probability of selecting a death move. This is in contrast with tone deletion for the TJM, where, once a death move has been selected, an existing tone is selected with uniform probability without considering the number of associated harmonics.

The next section will compare the performance of the different sampling schemes and the birth and death probabilities described here will be used to explain the behaviour of the algorithms.

6.3 Comparison of the Reversible Jump Schemes

This section compares the performance of the different reversible jump sampling schemes for tone and harmonic detection. The reduction in computational complexity is not quantified, however, it will be shown that the alternatives to the TJM derived here do not suffer any degradation in the quality of the estimates. The results first compare performance of the algorithms when the multiply/divide moves are not used by setting the parameter $\gamma = 1$. Subsequently, the impact of the multiply/divide move is illustrated by examining the performance of the CJM.

6.3.1 Overview of Experimental Setup

Test Signal 1**:** Two sets of experiments are considered in this work. The experiments are designed to highlight the effect of the sampling schemes used. The first experiment looks at the error associated with tracking a varying number of harmonics when a single tone with a

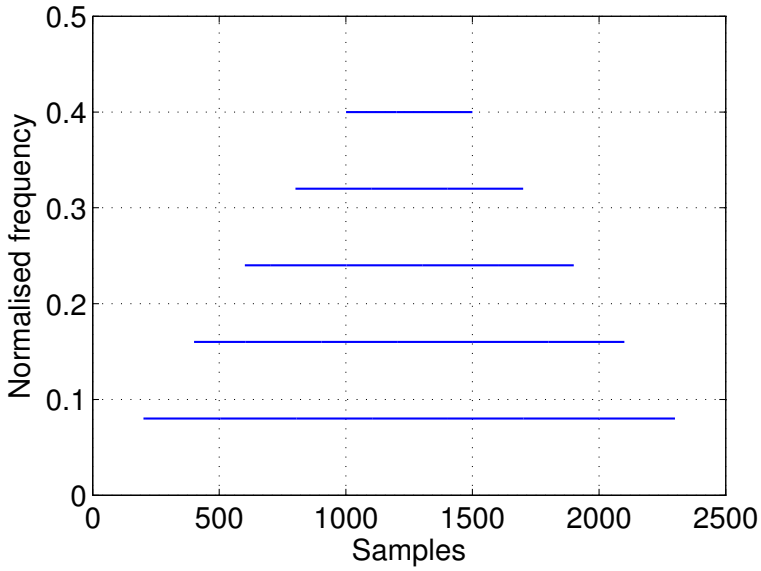


Figure 6.6: *Time-frequency representation of test signal* 1.

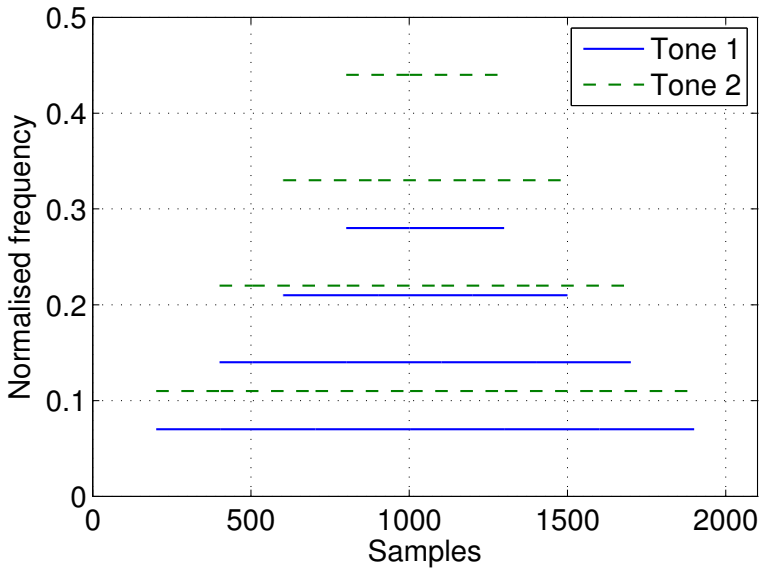


Figure 6.7: *Time-frequency representation of test signal* 2.

time-varying number of harmonics is present in the signal. The signal consists of a tone of normalised frequency 0.08 Hz, with an additional harmonic being added every 200 samples. After a duration of time, the number of harmonics is reduced by one every 200 samples as shown in the time-frequency representation in Figure 6.6.

Test Signal 2**:** The second test signal consists of two tones which follow a similar pattern of a first increasing and then decreasing number of harmonics as shown in Figure 6.7. In this case, the fundamental frequencies of the tones are 0.7 and 0.11 Hz. Since the number of harmonics changes simultaneously for both tones, this test signal is useful in examining the tracking-lag expected with the CJM as compared with the TJM and the MTJM.

Due to the use of a window in the algorithm, there is a mismatch with the model used when only a portion of the window overlaps a tone. At points of discontinuity in the signal, *i.e.*, when a tone/harmonic starts or stops, the algorithm may be expected to over-model the signal in an attempt to fit the discontinuity. This over-modelling appears as a manifestation of multiple frequency estimates in the vicinity of what is actually only a single tone or its harmonic. This over-modelling is common to all the reversible jump schemes, however, due to the nature of the construction of the reversible jump moves, some schemes are more prone to mis-modelling these discontinuities.

Experimental Setup: The different sampling schemes are simulated under identical setups. The number of particles used is 100 and the hyper-parameters for the truncated Poisson distributions are set as $\lambda_t = 1$ and $\lambda_h = 1$. Additionally, since a comparison is being made between the sampling schemes alone, the overtone deviation parameter μ, is constrained to zero in these simulations.The results from each algorithm are averaged over 100 runs for both test signals.

In addition to the TJM, MTJM and CJM, another scheme referred to as the single jump move (SJM) is also examined. The SJM scheme is identical to the TJM with the exception that the maximum number of harmonics is restricted, *i.e.*, $H_{max} = 1$, with the result that the second reversible jump for the number of harmonics is eliminated. The SJM thus approximates the signal as a set of independent tones without considering harmonics.

6.3.2 Estimation of a Single Tone with Multiple Harmonics

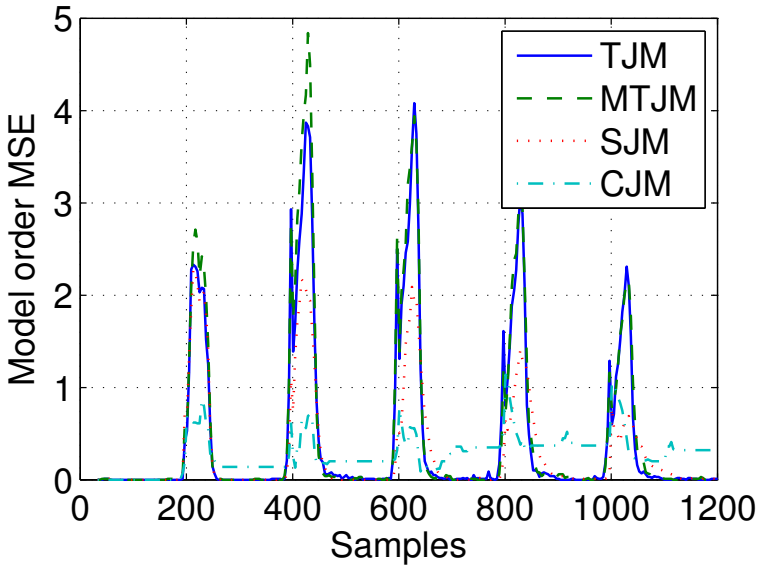
Figures 6.8 examines the extent of mis-modelling of the different algorithms when using *only* birth and death moves. The model order at each instant is defined as $O_k = \sum_{p=1}^{P_k} H_k^{(p)}$. Estimation of an excess number of tones or harmonics contributes to the error which is particularly seen when tones/harmonics start or stop.

The figure shows that all the algorithms tend to mis-model the signal at points of discontinuity in the signal. Additionally, the extent of mis-modelling at these points of discontinuities is seen to increase across the signal. This can be explained by examining the evolution of the δ^2 parameter. Figure 6.9 shows an example of the estimated value of δ^2 across the signal. Examination of the likelihood function in equation (4.40) shows that δ^2 acts as a penalty term. When δ^2 is large, addition of a new tone or harmonic must sufficiently offset the additional penalty incurred. The initial sampling distribution on δ^2 forces it to a large value, thus reducing the extent of mis-modelling. However, the value of δ^2 towards the end of the signal is significantly reduced, thus reducing the penalty from adding tones and harmonics. This results in further mis-modelling at points of discontinuity at the end of the signal as seen from the results.

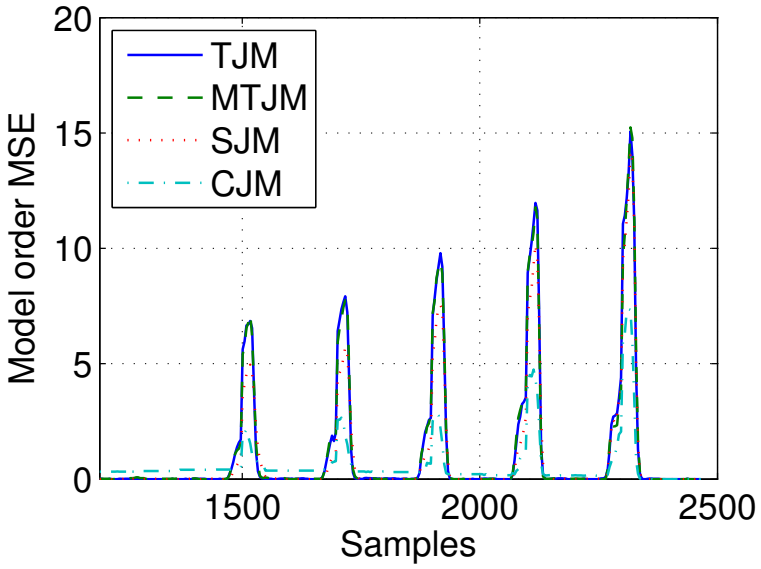
6.3.3 Estimation of Two Tones with Multiple Harmonics

To provide a quantitative comparison of the algorithms, the model order MSE of the different schemes for test signal 2 is listed in Table 6.1. It is seen that the performance of the SJM is comparable to that of the TJM and the MTJM. The model order MSE of the CJM is comparable with the other methods. However, the CJM detects the correct model order less often than the other methods as shown in Figure 6.8 as well as in Table 6.2. This decreased performance is a result of the algorithm getting caught in local maxima which will be discussed in the following section. For comparison, the results from incorporating the multiply/divide move in the CJM are included in the tables where it is seen that the performance of the algorithm is on par with the other methods, and in some cases, better.

From the figures, it is seen that the SJM does not mis-model to the same extent as the other algorithms. This can be explained by examining the probability of selecting a birth move as illustrated in Figure 6.1. Since the SJM assumes that harmonics are independent tones, as the number of harmonics increases, the SJM estimates an increasing number of tones. In the presence of a large number of tones, the probability of selecting a birth move decreases rapidly,



(a)



(b)

Figure 6.8: *Model order MSE for test signal* 1 *using only birth and death moves. The first half of the signal is shown in (a) and the second half is shown in (b).*

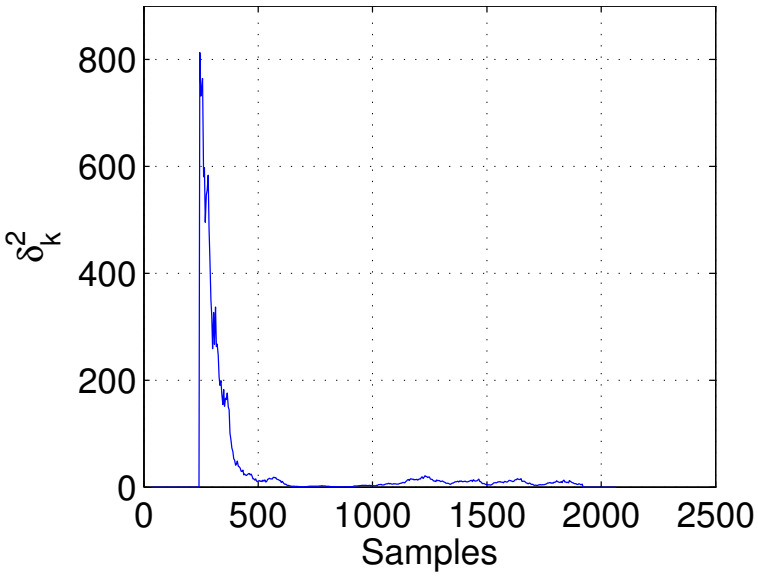
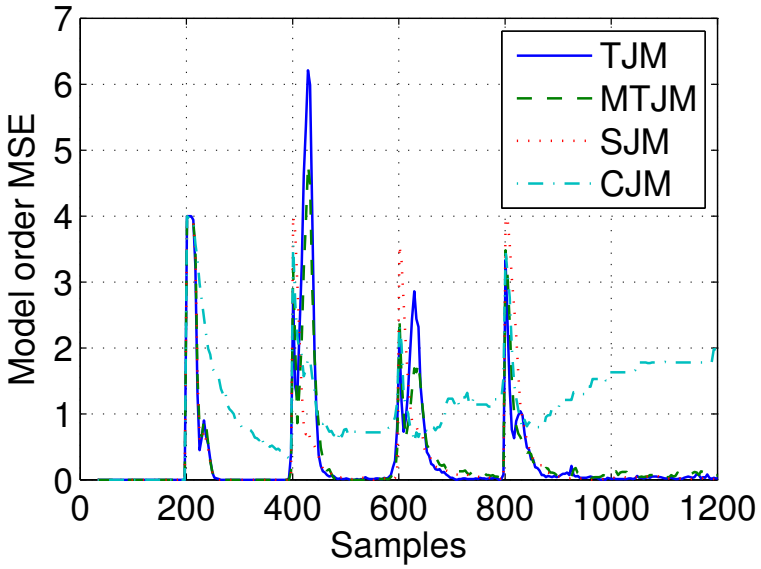


Figure 6.9: *Example of evolution of δ_k^2 over time for test signal* 2.

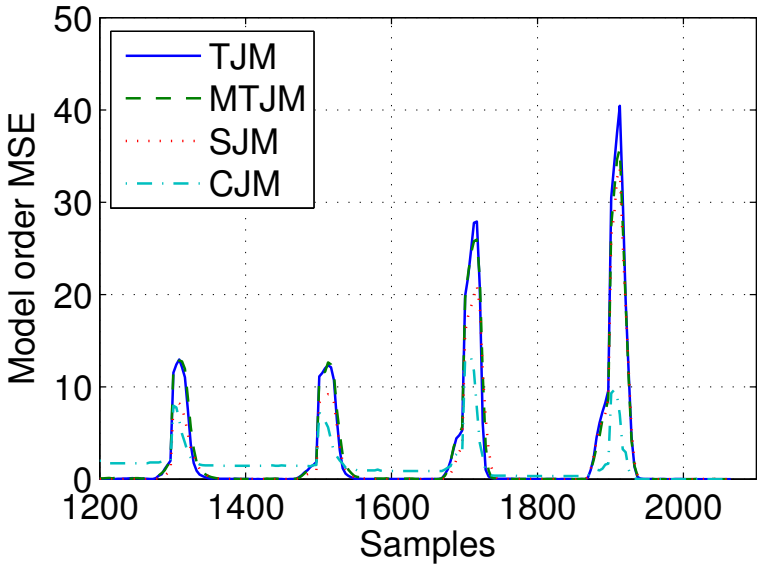
thus limiting the extent of mis-modelling. However, since the SJM does not assume harmonics in the signal structure, it is prone to larger MSE of the frequencies when dealing with such signals.

While the model order MSE of the SJM is comparable to the other methods, the MSE of the frequency estimates shows a marked increase when compared with the other methods, a direct consequence of treating the harmonics as independent tones. This is evaluated for the second test signal between samples 801 and 1300, when the largest number of tones and harmonics are present. During this segment of the signal, the percentage of estimates with the correct model order is obtained, and the frequency MSE is calculated using those points where the correct model order was estimated.

As seen in Table 6.2, the frequency MSE of the SJM is significantly larger since the harmonics are assumed as independent tones. The CJM is seen to perform poorly compared with the other methods. This is due to the lack of a multiply/divide move, and the tables show that incorporation of the multiple and divide moves result in significant improvements. The following section discusses the use of a multiply/divide move which will be shown to improve performance of the CJM.



(a)



(b)

Figure 6.10: *Model order MSE for test signal 2 using only birth and death moves. The first half of the signal is shown in (a) and the second half is shown in (b).*

Reversible jump scheme	Model Order MSE
TJM	1.3802
MTJM	1.3794
SJM	1.1624
CJM	1.3476
CJM*	1.3766

Table 6.1: *Comparison of model order MSE for test signal* 2 *for different sampling schemes. The* * *indicates the use of the multiply/divide move in the sampler.*

Reversible jump scheme	Model Order estimation accuracy (%)	Frequency MSE
TJM	88.5	0.0062
MTJM	86.6	0.0074
SJM	88.0	0.0107
CJM	81.2	0.0067
CJM*	90.0	0.0056

Table 6.2: *Comparison of model order estimation accuracy and frequency MSE for different sampling schemes applied to test signal* 2*. The* * *indicates the use of the multiply/divide move in the sampler.*

6.3.3.1 Importance of Multiply/Divide Moves

The multiply and divide moves are necessary to resolve ambiguities which may arise from erroneously tracking (sub-)harmonics rather than the fundamental tone. Figure 6.11 shows one such example where the filter gets caught in a local maxima when it tracks a sub-harmonic rather than the fundamental tone. The CJM is more susceptible to such behaviour since new tones are added along with harmonics, in contrast with the TJM and MTJM where a fundamental tone is first tested and then harmonics are added.

In this section, four minor variations of the CJM are considered. These variations arise from setting $\gamma = 0.75$ and $\gamma = 1$ which respectively enables and disables the use of the multiply/divide moves. Further, on addition of a new tone through a birth move, the number of harmonics can be either restricted to $n_b = 1$, or variable and sampled according to equation (5.26) (denoted

by 'var' in the table of comparisons). Figure 6.12 and Table 6.3 compare the performance of the CJM under these different scenarios. The table illustrates the model order MSE as well as the percentage of time for which the correct model order is determined.

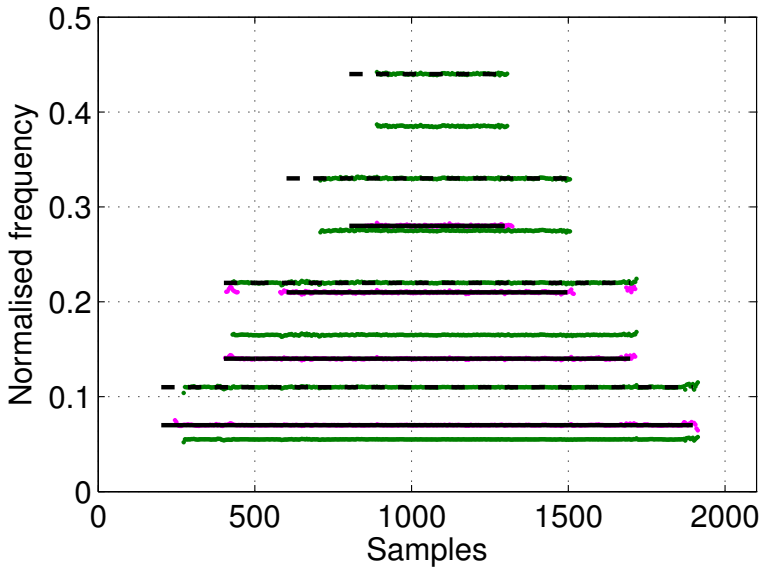
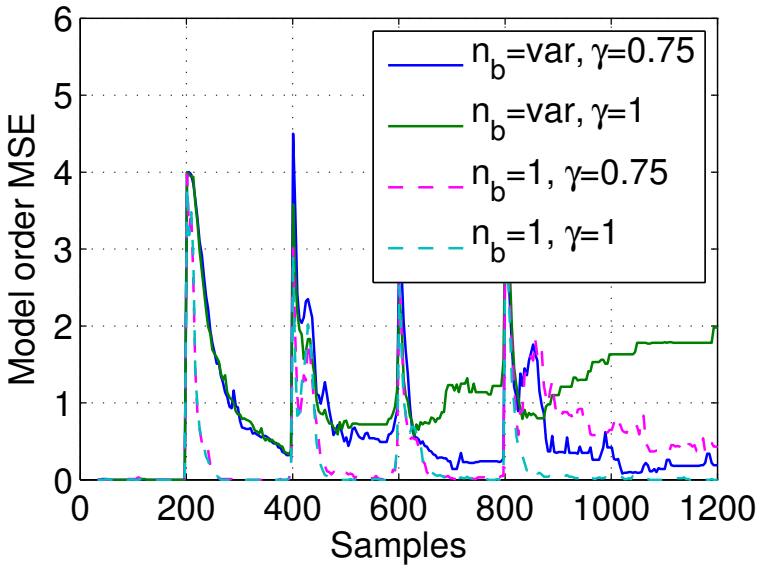


Figure 6.11: *Example showing the CJM getting caught in a local maxima and tracking a sub-harmonic component.*

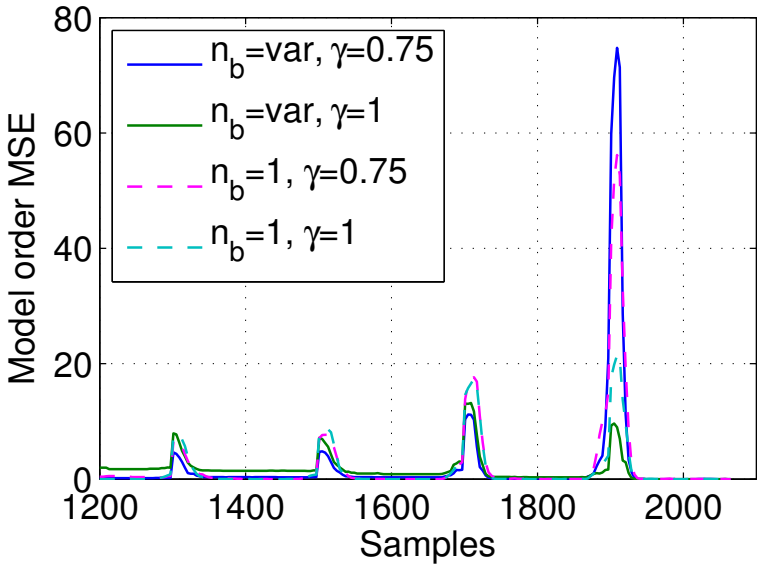
Case	n_b	γ	Model order MSE	Model Order estimation accuracy (%)
1	var	1	1.3476	73.7
2	var	0.75	1.3766	79.4
3	1	0.75	1.3198	81.3
4	1	1	0.8002	83.6

Table 6.3: *Comparison of CJM performance under different settings for test signal* 2*.*

Case 1 **and** 2**:** Cases 1 and 2 compare the performance of the CJM without, and then with the use of the multiply/divide move. It is seen that inhibition of the multiply/divide move results in increased modelling error as seen in Case 1. The addition of the multiply/divide move results in a significant increase in the percentage of correct model order estimates since the algorithm no longer gets caught in the local maxima traps.



(a)



(b)

Figure 6.12: *Model order MSE for the CJM under different conditions for test signal 2. The first half of the signal is shown in (a) and the second half is shown in (b).*

Case 3**:** Case 3 is a variation of case 2 with n_b constrained to 1. From Figure 6.12, it can be seen that under these constraints, the algorithm is seen to converge more quickly compared to either of the previous cases. This is most likely due to a better directed search of the parameter space in the reversible jump move, since in fact, multiple harmonics are not being added simultaneously, and this is reflected by the constraint $n_b = 1$.

Case 4**:** Case 4 shows an interesting deviation from what may have been the expected trend in the results. In this case, n_b is limited to one, and only a birth/death move is used for detection of the harmonics. The results can be explained by considering the structure of the test signal used – the harmonics in the signal appear sequentially, with the fundamental tone appearing first. Thus, the algorithm need only detect the fundamental tone prior to the appearance of the harmonics and this is easily accomplished using only the birth and death moves. Additionally, since n_b is constrained to one, the possibility of ambiguities (through the tracking of sub-harmonics) is severely diminished as seen in the previous case. Consequently, particles which would otherwise be devoted to the multiply/divide move can be used to track the birth and death of the harmonics thus leading to faster convergence and lower errors.

6.3.4 Conclusions

The above presented results have provided a comparison between three different sampling schemes. Comparing the different levels of performance, it is seen that the CJM offers an attractive alternative to using the TJM. The CJM performs no worse than the TJM for the test signals used, whilst reducing the computational complexity of the algorithm. Additionally, the assumption of a slowly varying signal does not significantly impact the performance as seen from the results.

6.4 Estimation of Inharmonicity

The general harmonic model (GHM) accounts for deviation in the harmonics from exact integer multiples. The need for modelling inharmonicity arises when dealing with certain musical instruments [38]. The time-varying nature of the model used here allows the inharmonicity parameters μ to vary over time, although this is not strictly necessary. In cases where the inharmonicity is a static parameter, alternate approaches may be required (see for example,

[95]).

Detection of inharmonicity in bat echolocation calls is an aspect of interest to biologists in the field. Before applying the algorithm to such calls, it is necessary to establish the accuracy of the estimate of the overtone deviation parameter μ. A significant deviation in overtones is specified as one that exceeds 0.01 (or 1%). The accuracy of the estimates is thus measured using a test signal possessing a deviation of 1%.

6.4.1 Estimation of Inharmonicity in an FM Signal

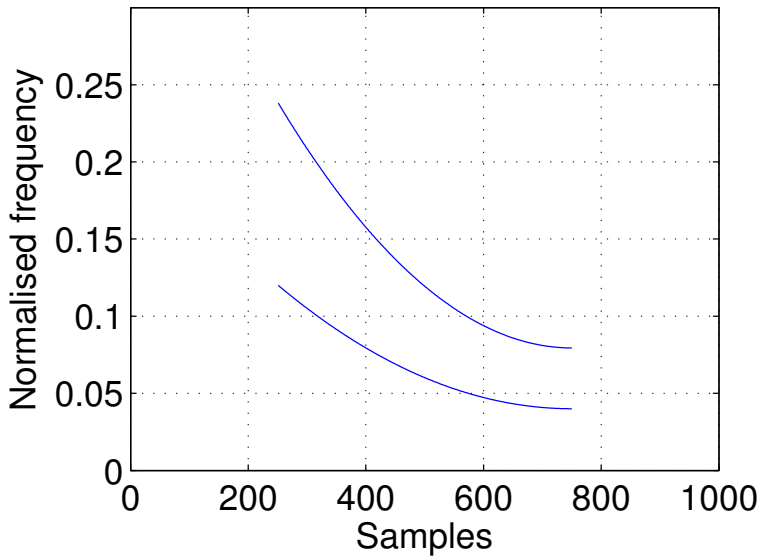


Figure 6.13: *Time-frequency representation of the test signal.*

The synthetic test signal consists of a quadratic chirp with an additional harmonic, such that the (fundamental) frequency sweeps from 0.12 to 0.04 Hz over 1000 samples as shown in the time-frequency representation in Figure 6.13.

The algorithm is constrained to detect a single tone with a variable number of harmonics. Additionally, the prior on the deviation is specified as $p(\mu) = \mathcal{U}[-0.025, +0.025]$ which allows a maximum deviation of 2.5%. Simulations are run using different numbers of particles, and the results are averaged over 100 independent runs.

Figure 6.14 shows the mean and standard deviation of the estimates averaged over 100 runs when using 100, 250, 500 and 1000 particles. The average standard deviation from each case

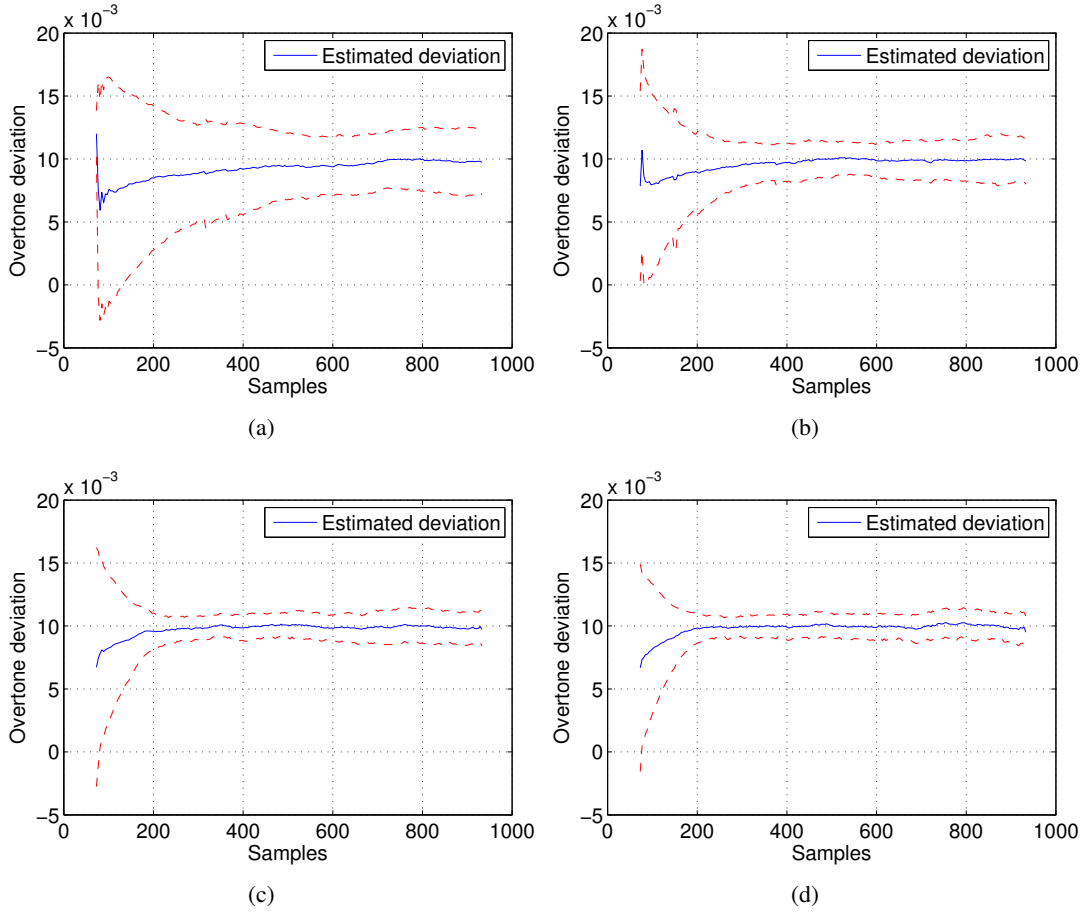


Figure 6.14: *Mean and standard deviation of overtone deviation parameter using* (a) 100*,* (b) 250*,* (c) 500 *and* (d) 1000 *particles.*

is compared in Figure 6.15. From the figure, it is seen that when using 100 particles, the algorithm is slow to converge to the true value of overtone deviation. It can also be seen that the gain from doubling the number of particles from 500 to 1000 produces only a marginal improvement. Table 6.4 shows the mean value and standard deviation averaged over the entire signal. Additionally, these values are also calculated by ignoring estimates from the first 200 samples and using only the following stable region, providing an asymptotic result.

The results show that the algorithm is capable of estimating an overtone deviation of 1% with a standard deviation of 0.001 when using 1000 particles. This level of accuracy is acceptable for the analysis of bat echolocation calls.

Number of particles	Mean $\pm$ standard deviation (10^{-3})	Mean $\pm$ standard deviation† (10^{-3})
100	9.19 ± 3.67	9.42 ± 3.02
250	9.60 ± 2.25	9.79 ± 1.68
500	9.73 ± 1.58	9.90 ± 1.14
1000	9.79 ± 1.39	9.97 ± 1.02
1000‡	10.00 ± 1.67	10.01 ± 1.05

Table 6.4: *Mean and standard deviation of the overtone deviation parameter averaged over the entire signal and* †*ignoring the first* 200 *samples of the signal. Mean and standard deviation for the amplitude modulation test signal are indicated by* ‡.

6.4.2 Estimation of Inharmonicity in an AM-FM Signal

The test signal used in the above comparison is only frequency modulated. Since bat echolocation calls are also amplitude modulated, it is necessary to establish the performance under these conditions. The signal model assumes that the amplitude is constant within the window, however, this assumption may not always hold true. The following discussion looks at the case when the test signal contains amplitude modulation as well.

In order to create the amplitude modulated test signal, the envelope from an actual bat call is used to modulate the test signal. Figure 6.16 shows the envelope used to modulate the signal and the resulting AM-FM time-domain signal. This test signal violates the assumption of constant amplitude within the window and serves to verify the accuracy of overtone deviation estimates when such is the case.

As in the previous experimental setup, the algorithm is identically constrained and the results are averaged over 100 runs using 1000 particles. The mean and standard deviation averaged over all the runs is shown in Figure 6.17. Table 6.4 compares the results from the amplitude modulated test signal with those from the unmodulated signal. From the table, it is seen that there is a slight increase in the standard deviation of the estimates, which are otherwise accurate.

The performance of the algorithm is comparable to that of the unmodulated signal discussed previously, and it is clear that the algorithm can be used to detect deviations in overtones which exceed 1%. Based on these results, Section 7.3 in the following chapter examines the presence of significant levels of inharmonicity which may occur in bat echolocation signals.

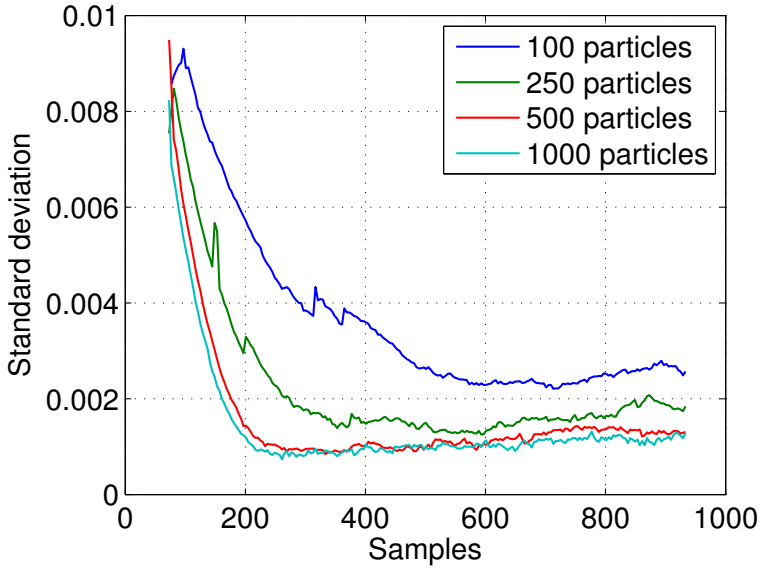


Figure 6.15: *Comparison of the standard deviation of the overtone deviation estimate when using a varying number of particles.*

6.5 Summary

This chapter examined alternatives to the TJM described in the previous chapter as well as the estimation of inharmonicity in the overtones. The TJM involves the application of a first reversible jump move to determine the number of tones, followed by a second reversible jump move, applied to each tone, to determine the number of associated harmonics. The application of each reversible jump move involves evaluation of an acceptance probability which leads to a high degree of computational complexity.

It is possible to reduce the complexity of the moves by incorporating simplifying assumptions which limit the range of possible moves. The main assumption is that the signal varies slowly, such that only a single tone can undergo any change at any instant of time. This leads to the first of two alternatives called the MTJM where in addition to the first reversible jump move for the number of tones, a single reversible jump is applied to a single tone selected at random to determine the number of associated harmonics.

The second reversible jump scheme, the CJM involves combining the two moves yielding a single reversible jump move which is applied to each particle. Again, assuming that only a single tone can undergo a change at any instant, birth and death moves involve addition or removal of n harmonics of an existing or new tone.

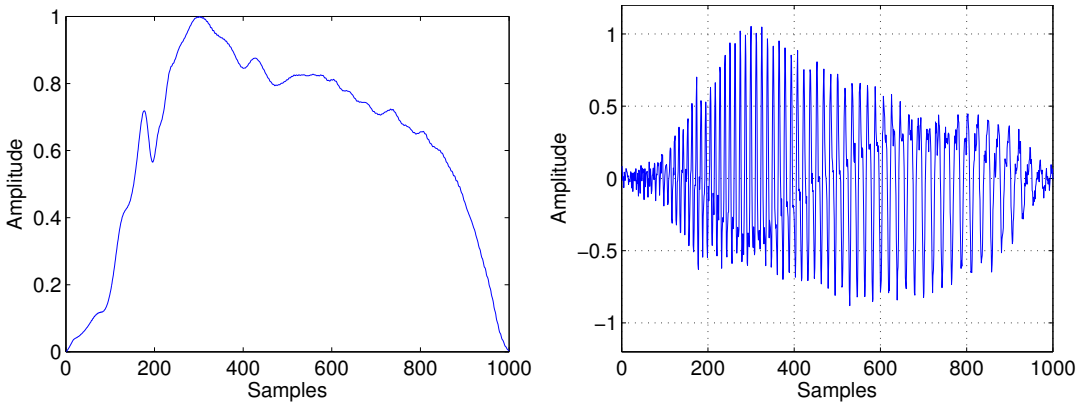


Figure 6.16: *a) Amplitude envelope of a bat call used to modulate the test signal and, b) the resulting time-domain AM-FM signal.*

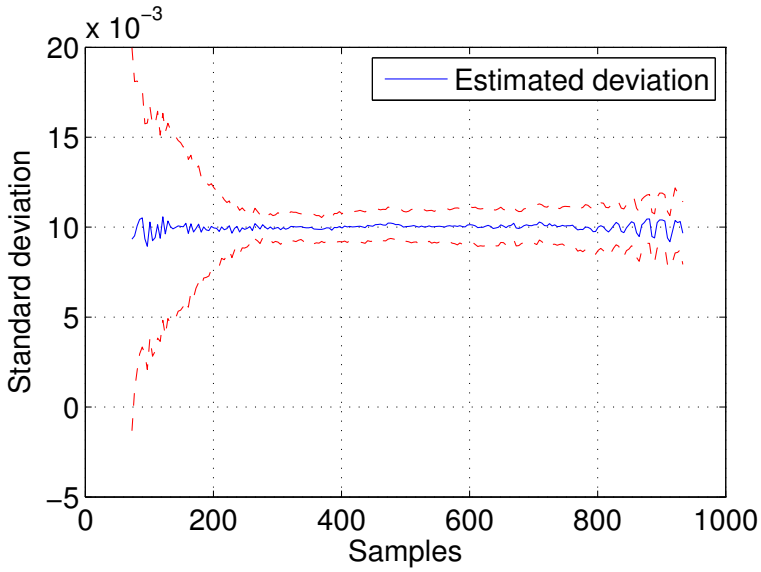


Figure 6.17: *Mean and standard deviation of the overtone deviation parameter for the amplitude modulated test signal.*

Section 6.2 discusses the transition probabilities of birth and death moves for the different algorithms. In the case of the TJM and the MTJM, the birth and death probabilities for the number of harmonics is independent of the number of tones present. The CJM, in contrast incorporates a dependence such that it becomes possible to specify a region of preference specified by both the number of tones and their harmonics.

Section 6.3 compares the performance of the different reversible jump schemes and shows that the new algorithms developed in this chapter work as well as the original TJM but operate with reduced computational load.

The chapter concludes with a discussion of the estimation of inharmonicity in overtones. Comparisons are performed using synthetic data and a varying number of particles. The results show that the algorithm is capable of detecting a deviation of 1% with reasonable accuracy. The results also indicate that the algorithm is capable of estimating the deviation in the presence of amplitude modulation. This is of particular importance since bat echolocation calls can be highly amplitude and frequency modulated.

The following chapter looks more closely at application of the sequential frequency estimation algorithm to bat echolocation calls. The chapter examines selection of the polynomial basis as well as the presence of inharmonicity in calls.

Chapter 7
Analysis of Bat Echolocation Calls

Introduction

This chapter examines aspects of performance related to the use of polynomial chirps as the basis in the signal model proposed in Chapter 4 and in particular, benefits from their application to bat echolocation calls. The previous chapters have used a sinusoidal (zeroth-order polynomial) basis for all comparisons. The aim of this chapter is to illustrate how higher order polynomials can provide a better fit under certain circumstances. The use of higher order chirps, however, comes at a cost, since additional chirp parameters are required to be estimated. This can limit the usability of higher order polynomials, particularly when a limited number of particles is used in the filter. The tradeoffs between the basis flexibility and model simplicity will be discussed in Section 7.1 before continuing with application of the algorithm to the study of bat echolocation calls.

Section 7.2 considers the application of the sequential frequency estimation algorithm to bat echolocation calls. Aspects include the relevance of a higher order polynomial basis and the impact of window length on the frequency estimates. This is followed by an examination of inharmonicity in bat echolocation calls. The previous chapter presents results which impose limits on the accuracy with which inharmonicity can be measured using the algorithm developed here, and the presence of inharmonicity is examined within these constraints.

The concluding section of the analysis of bat echolocation calls looks at *feeding buzz* sequences. The feeding buzz, described in more detail in Section 7.4, consists of extremely short duration calls which makes it difficult to obtain an accurate time-frequency representation of these signals. Using the method developed here, it will be shown that an accurate time-frequency representation of these calls can be obtained, thus facilitating further study of these calls.

7.1 Choice of Basis Polynomial Order

Selecting the order of the polynomial basis is not a straightforward matter since it depends on the desired estimation accuracy (and hence, the window length) and the structure of the signal under consideration. This section aims to shed some light on how such a selection may be made by considering the MSE arising from the use of different polynomial orders.

7.1.1 Effect of Polynomial Basis Order on Frequency MSE with Constrained Model Order

The model used for sequential frequency estimation defined in Chapter 4 describes how the signal is analysed within a sliding window. The variance on the estimated frequency reduces as the length of the window increases. Increasing the window length, however, may result in increased mismatch between the windowed signal and the basis function. To illustrate this trade-off, a quadratic chirp signal is analysed using both a zeroth- and a first-order polynomial basis with varying window lengths.

Figure 7.1 shows the instantaneous frequency of the test signal, which consists of a quadratic chirp with a CF tail. Figure 7.2 shows a high contrast spectrogram of the test signal (without noise) using a window length of 129 samples. Using this window length, multiple peaks in the spectrogram are visible. The presence of these multiple modes is due to the large bandwidth of the signal within the window. Under these circumstances, the use of a zeroth-order polynomial basis will result in over-modelling as the algorithm attempts to fit multiple sinusoids to the large bandwidth chirp within the window.

By way of comparison, the performance of both basis functions is examined at an SNR of 12 dB. In the first comparison, the algorithm is constrained to estimation of a single tone, *i.e.*, $P_{min} = P_{max} = 1$, $H_{min} = H_{max} = 1$, and the frequency MSE using both basis functions is averaged over 200 independent runs using 100 particles.

The resulting time-varying frequency MSE is illustrated in Figure 7.3 for varying window lengths. For short window lengths, there is almost no difference in performance between the zeroth- and first-order basis functions. As the length of the window increases, the difference in performance between the basis functions becomes more apparent. As the window length is increased further, the zeroth-order basis performs better in the CF portion of the signal and worse

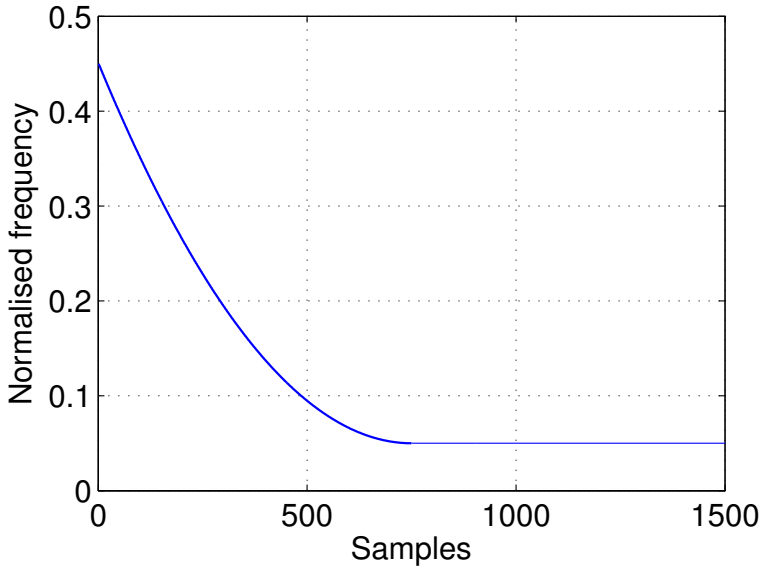


Figure 7.1: *Instantaneous frequency of quadratic chirp test signal.*

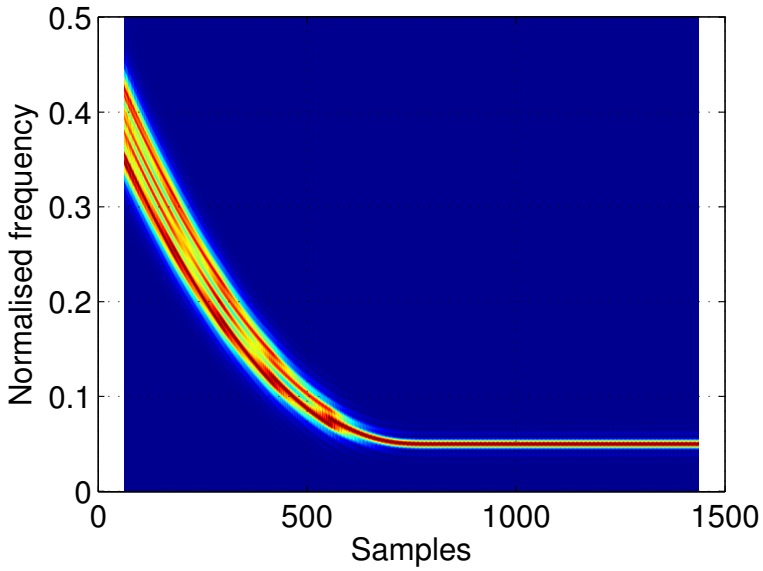


Figure 7.2: *Spectrogram of quadratic chirp test signal. Multiple ridges are present in the quadratic chirp segment which are responsible for over-modelling when using a purely sinusoidal model.*

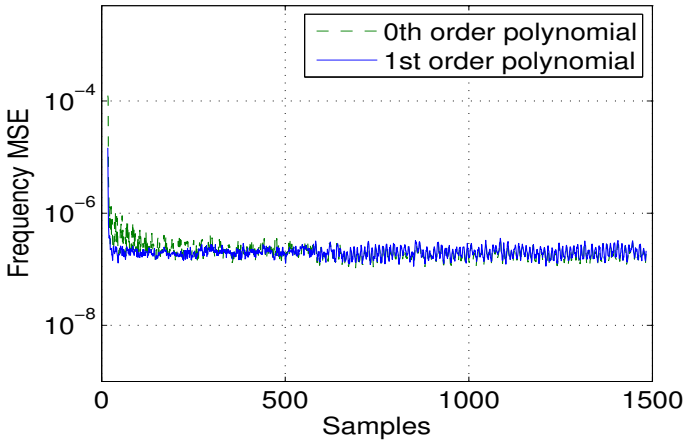
in the FM portion. The linear chirp basis, in contrast, performs no better for the CF portion, and in fact, starts to show a higher frequency MSE in the FM portion of the signal. This effect is due to the limitation of the linear chirp which is unable to characterise the frequency modulation in the 129 sample window adequately. Using a quadratic chirp basis would alleviate this constraint, however, a larger number of particles would then be required to achieve a similar error performance. This will be illustrated later in Section 7.1.3.

From the results, it is seen that the linear chirp basis offers better performance when the signal in question is frequency modulated. The difference in performance between the sinusoidal and linear chirp basis functions decreases with a decrease in the window length, although, the overall frequency MSE may increase due to the presence of less data in the window.

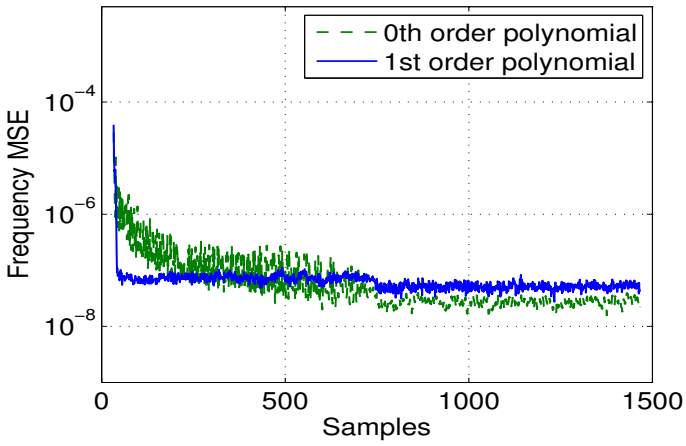
7.1.2 Effect of Polynomial Basis Order on Frequency and Model Order MSE with Unconstrained Model Order

A second comparison is made between the model order MSE using the zeroth- and first-order polynomial basis functions, where the model order is defined as $O_k = \sum_{p=1}^{P_k} H_k^{(p)}$. In this comparison, only the number of harmonics is constrained $H_{min} = H_{max} = 1$, while the number of tones P_k is unconstrained. Figure 7.4 illustrates the model order MSE for the two basis functions for window lengths of 65 and 129 samples. Similar to the previous results, it is seen that the zeroth-order polynomial performs poorly, except in the CF segment of the test signal. As mentioned earlier, this happens due to over-modelling of the high bandwidth chirp using multiple sinusoids.

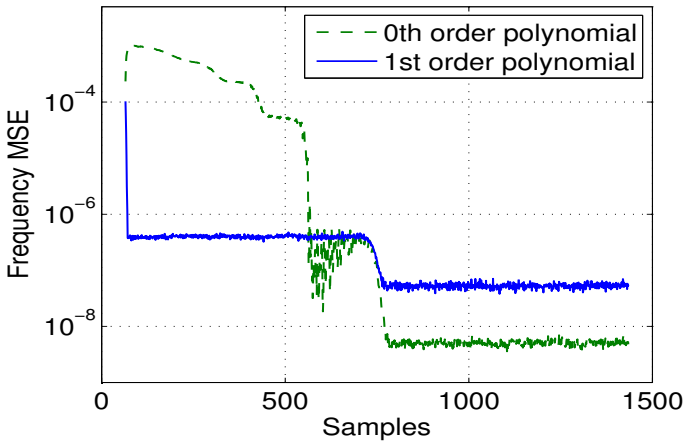
Figure 7.5 compares the frequency MSE for polynomial orders zero through four. The frequency MSE is evaluated only at instants where the correct model order has been determined. An example run of the zeroth-order polynomial is shown in Figure 7.6 which can be used to explain the frequency MSE curve. Due to the large bandwidth of the chirp in the initial stages, the algorithm prefers to model the signal using a noise only model, while in latter parts of the frequency modulated segment, the algorithm fits multiple tones to the signal, resulting in erroneous estimates of the number of tones. Consequently, the zero order polynomial basis is on average able to detect the correct model order only 47.6% of the time, in contrast with higher order polynomials which detect the correct model order over 90% of the time as illustrated in Table 7.1.



(a)

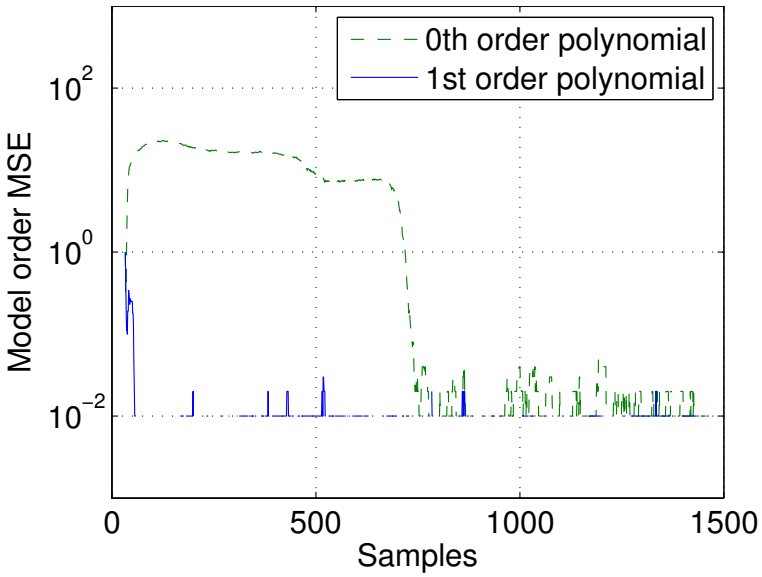


(b)

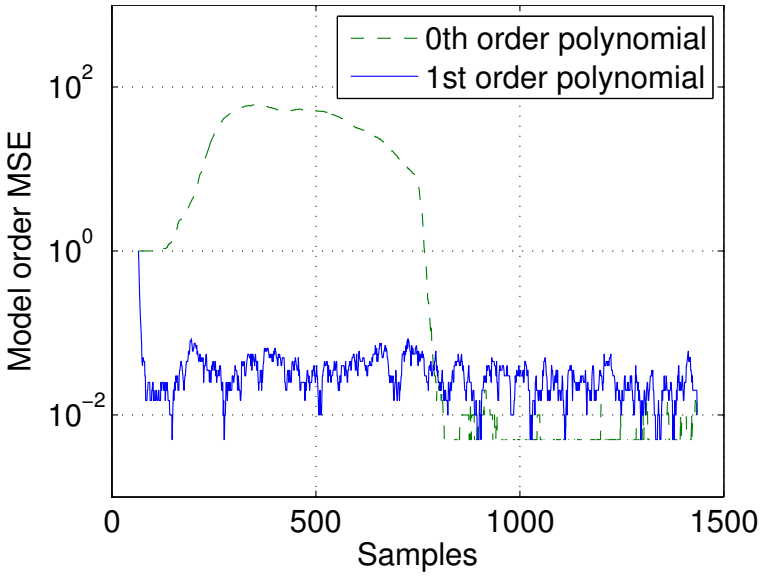


(c)

Figure 7.3: *The frequency MSE is compared for zeroth and first-order polynomial basis functions using window lengths of (a)* 33*, (b)* 65 *and (c)* 129 *samples.*



(a)



(b)

Figure 7.4: *The model order MSE is compared for zeroth and first-order polynomial basis functions using window lengths of (a)* 65 *and (b)* 129 *samples. Large signal bandwidth causes incorrect model order estimation for the zero-order polynomial basis.*

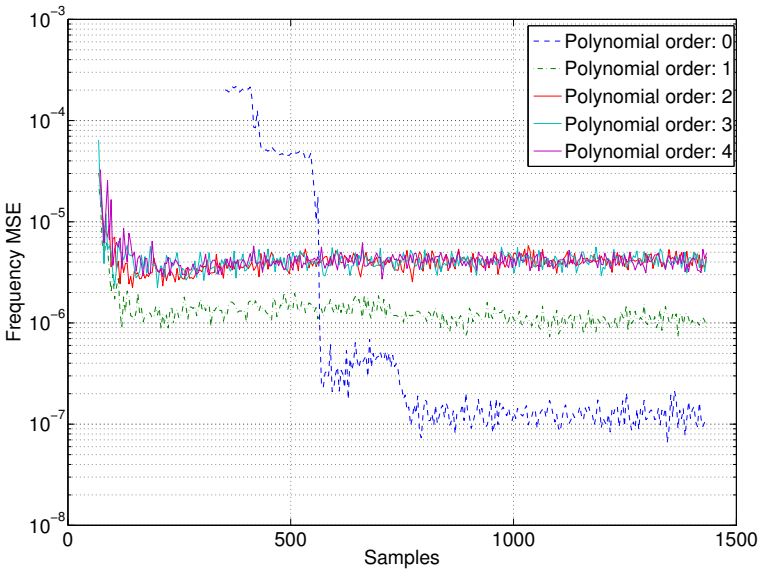


Figure 7.5: *Comparison of frequency MSE for varying polynomial basis orders.*

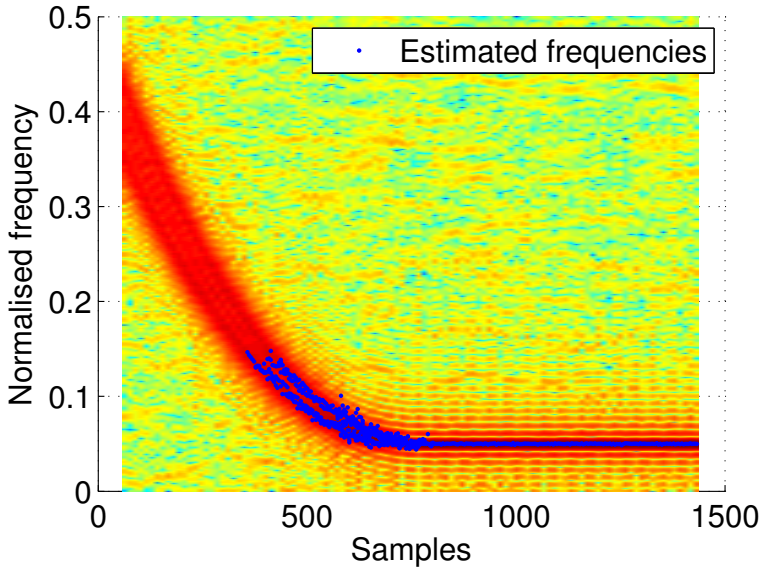


Figure 7.6: *Example of estimated frequency content of a quadratic chirp using a zeroth-order polynomial basis.*

Polynomial Order	Model Order estimation accuracy (%)	Frequency MSE (10^{-6})
0	47.6	22.00
1	93.2	1.44
2	93.2	4.17
3	93.0	4.33
4	93.7	4.43

Table 7.1: *Comparison of model order estimation accuracy and frequency MSE for different polynomial orders.*

As illustrated in both Figure 7.5 and Table 7.1, in contrast with the poor performance of the zeroth-order polynomial, a first-order polynomial provides a near constant frequency MSE throughout the duration of the signal. Also of interest is the performance of the higher polynomial order models which appear to provide a ceiling on the error, with the frequency MSE remaining constant with increased polynomial order. Thus, with the lowest frequency MSE, the first-order polynomial may be considered to be the optimal basis for the quadratic chirp test signal for the window lengths considered here.

7.1.3 Error Performance with Varying Number of Particles

The previous results are obtained using an identical number of particles for each model. Under circumstances where a more accurate representation of the signal is desired, a quadratic or higher order polynomial may be required. Figure 7.7 compares the performance of a zeroth-, first- and second-order polynomial chirp basis using 100 and 1000 particles. It is seen that the performance of the linear chirp basis using 1000 particles approaches the performance of the sinusoidal basis which uses 100 particles. At the same time, using ten times as many particles, the quadratic chirp basis performs as well as the linear chirp basis which uses 100 particles. Any improvements in performance, however, are constrained by the ability of the basis to accurately represent the frequency modulation in the signal. Increasing the number of particles does not reduce the mismatch between model and observations and thus does not always improve performance.

The quadratic chirp test signal used here is not representative of the large class of signals that

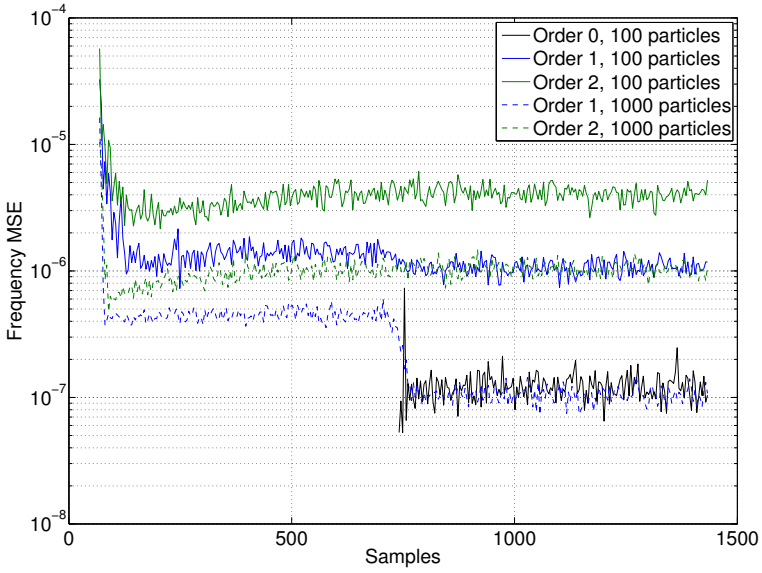


Figure 7.7: *Comparison of frequency MSE for different polynomial orders and number of particles.*

may need to be examined. The comparison using this signal, however, serves to illustrate the trade-off between model flexibility and variance of the estimates for a given number of particles. The next section will apply the sequential frequency estimation algorithm to bat echolocation calls. The relevance of higher order chirp basis functions to characterise detail in the FM signals, inharmonicity in the echolocation calls, and feeding buzz sequences will be discussed.

7.2 Relevance of the Polynomial Chirp Basis to the Analysis of Bat Echolocation Calls

The previous section considered the application of different polynomial chirp basis functions to a synthetic signal. The synthetic signal does not contain the variation which is found in bat echolocation calls. This section addresses the choice of polynomial chirp order and window length in the context of analysis of bat calls. These two parameters are linked, in that the use of a longer window length may require a more flexible basis, which in turn may require the use of a larger number of particles in order to maintain the same error performance. In general, a longer window is desirable since the variance on the resulting frequency estimates will be lower. Two bat echolocation calls are chosen for this comparison. The first call is from *P. pipistrellus* while the second is from *M. daubentonii*.

Since the true instantaneous frequency of the signal is unknown, an error term cannot be used as criteria for comparison. Instead, the variance of the distribution on the instantaneous frequency is used to compare the different chirp basis functions. In this case, the variance can be viewed as a measure of how well the model fits the observations. If a model with excessive flexibility, *i.e.*, high polynomial order, is used, then it may be possible to fit the observations using a set of parameters in which the frequency deviates further from the true frequency value. In comparison, due to the lack of flexibility, a low polynomial order will be more sensitive to deviations from the true parameters. In this way, the variance on the frequency estimate is used as a measure of the goodness of fit of the basis.

In addition to examining the variance of the frequency estimates, the residuals and spectrograms of the residuals are also examined. The residual at time t, ϵ_t is computed as:

$$\epsilon_t = x_t - \hat{\mathbf{G}}_t \hat{\mathbf{a}}_t \tag{7.1}$$

where x_t is the observation, the matrix $\hat{\mathbf{G}}_t$ is constructed as a function of the estimated parameters $\hat{\boldsymbol{\psi}}_t$, and $\hat{\mathbf{a}}_t$ is the estimated amplitudes as given by equation (5.19). If the residuals possess a noise-like structure, then the algorithm may be judged to have captured all the information of the signal. The residuals will be used to compare the performance of the MPF- and RBPF-based algorithms in the following sections.

A large number of particles is used to characterise the detail in the posterior distribution, with the result that the Monte Carlo estimate approaches the true posterior distribution. Consequently, the variance of the parameter estimates obtained from the filter can be regarded as the variance of the true posterior distribution. The algorithm is constrained to detecting a *single* tone, but with a *variable* number of harmonics and the number of particles is set to $10{,}000$, such that the posterior distribution is sufficiently well sampled. The algorithm is tested using polynomial basis functions of orders one through three, as well as with window lengths of 65 and 129 samples, which corresponds to 0.087 and 0.172 ms respectively, since the signals are recorded at a sampling frequency of 750 kHz. Additionally, inharmonicity is not considered in these simulations.

Figure 7.8 shows a recording of a pipistrelle call along with the spectrogram. The time domain signal shows the presence of amplitude modulation, while the spectrogram illustrates the frequency modulation and indicates the presence of overtones in the signal. The second echolo-

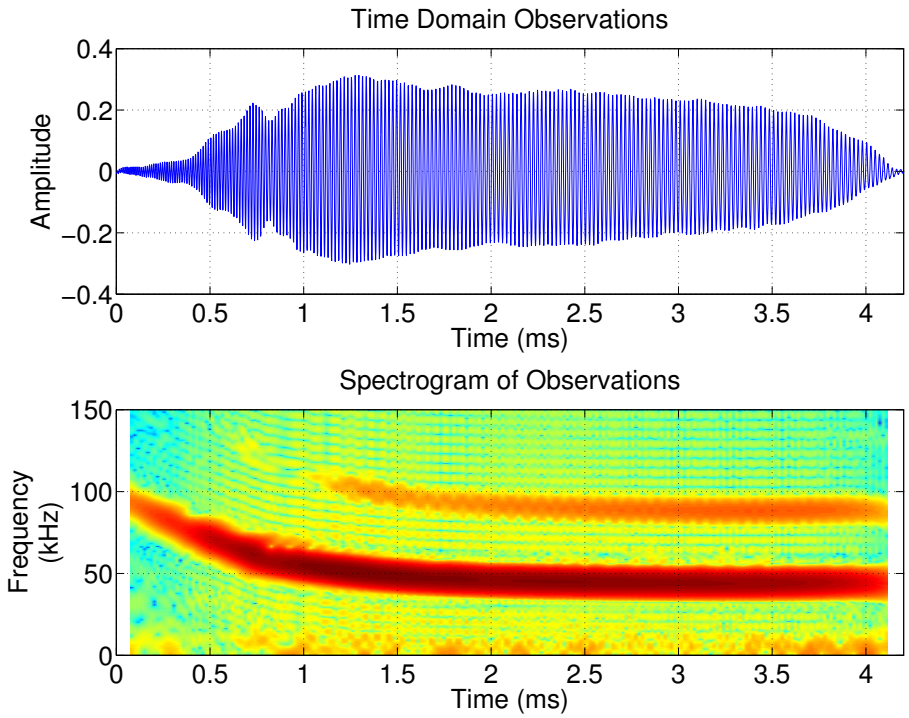


Figure 7.8: *Time domain and time-frequency representation of a call from* P. pipistrellus.

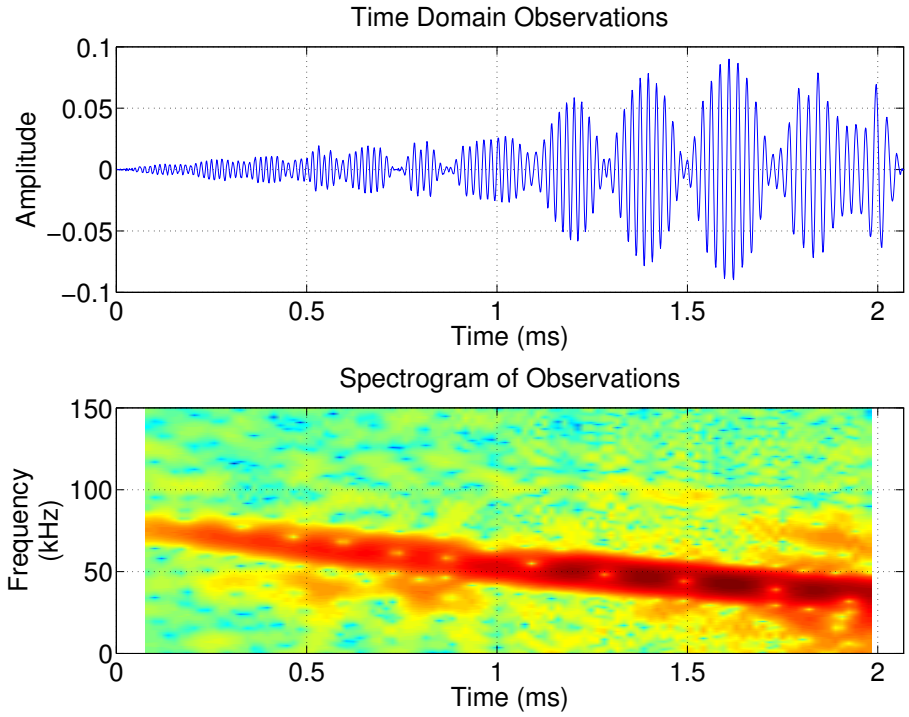


Figure 7.9: *Time domain and time-frequency representation of a call from* M. daubentonii.

cation call used for analysis is shown in Figure 7.9. The call from Daubenton's bat does not possess additional overtones, however, it is seen to possess notches at regular intervals which manifest as nulls in the spectrogram. The following discussion will illustrate how polynomials of different orders perform with these signals.

7.2.1 Analysis of Call from *P. pipistrellus*

Figure 7.10 illustrates the estimated instantaneous frequency for the different polynomial orders using window lengths of 65 and 129 samples. In the case of the 65 sample window, the performance of the different basis functions appears similar. Due to the use of a short window, the additional polynomial terms do not provide sufficient benefit to outweigh the cost of estimating those additional terms. The impact of the additional polynomial terms is more visible when a window length of 129 samples is used. In the case of the longer window, the impact of the higher polynomial terms will be larger further away from the centre of the window, so that minor deviations in the values of these parameters can have a large effect on the posterior distribution. Additionally, any such deviations get multiplied for overtones, further increasing the mismatch between the model and the observations.

It is thus seen that the higher order polynomials are unable to track the overtone as well as a first-order polynomial basis. For the pipistrelle call considered here, while the underlying spectrogram has a significant spread of energy in the time-frequency plane, the sequential frequency estimator is able to obtain a clear point estimate of the time-varying frequencies.

As mentioned earlier, the variance on the frequency estimate is used as a means of comparison between the different polynomial basis functions. Figure 7.11 compares the variance of the estimated frequency for the first 500 samples of the signal for a window length of 65 samples. It can be seen that the first-order polynomial basis outperforms the higher order polynomials, producing a lower variance on the frequency estimates. Figure 7.12 compares the performance of the first-order polynomial basis for window lengths of 65 and 129 samples, where it is seen that the use of the longer window decreases the variance on the frequency estimate.

Following analysis of the variance on the frequency estimates, the residuals are examined next. Figure 7.13 shows the residuals and the spectrogram of the residuals for a window length of 65 samples for polynomial orders 1 through 3. The spectrograms are used to illustrate that the residuals are not completely noise-like and harmonic components are seen in the spectrograms.

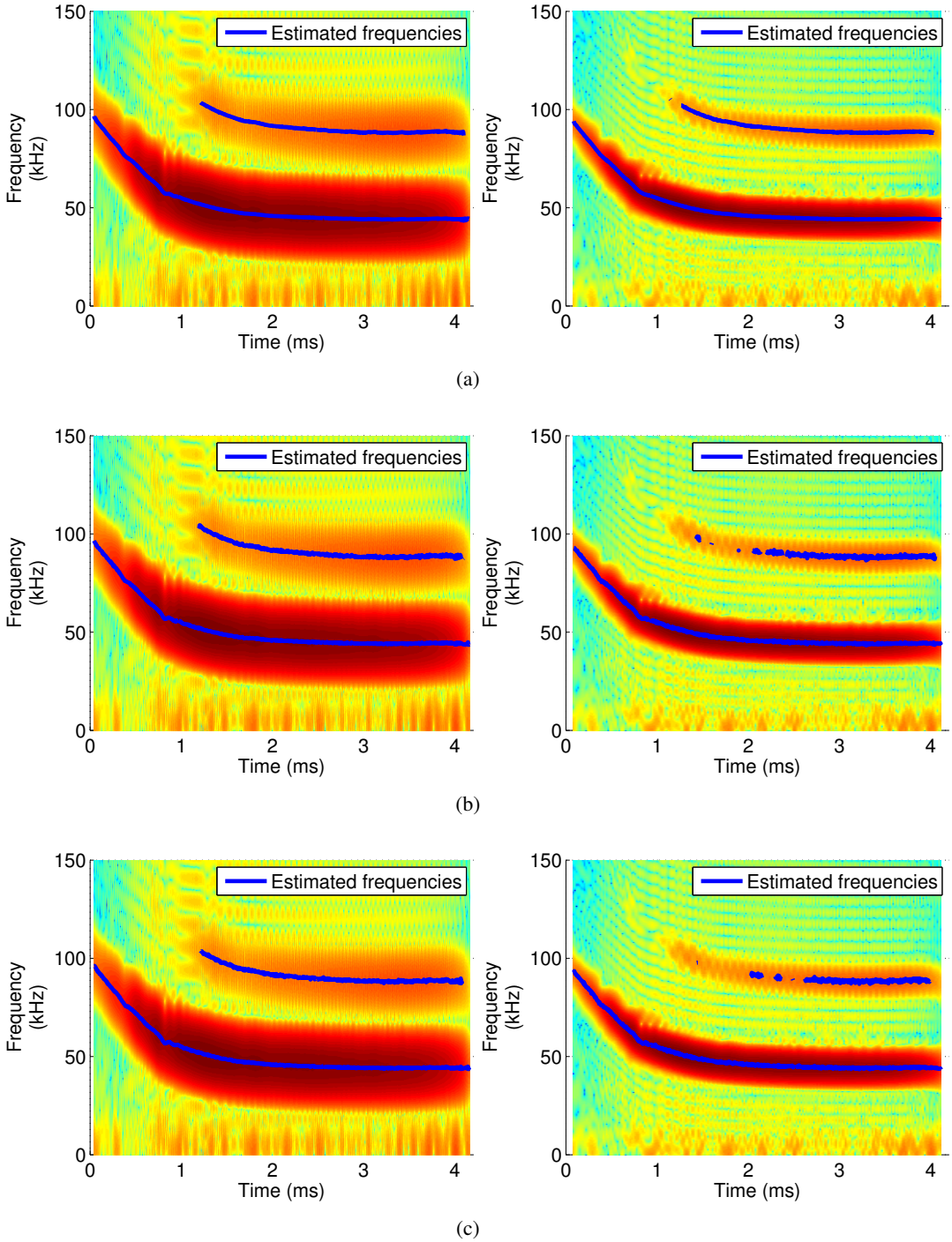


Figure 7.10: *Instantaneous frequency estimates for the call from* P. pipistrellus *using the MPF using polynomial basis orders (a)* 1*, (b)* 2 *and (c)* 3*. The figures on the left demonstrate results using a window length of* 65 *samples and those on the right use a window length of* 129 *samples.*

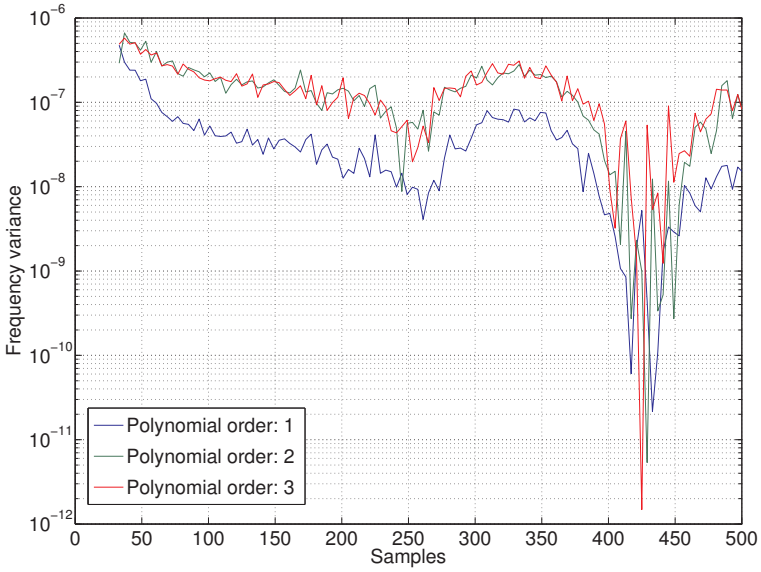


Figure 7.11: *Variance on the frequency estimate for varying polynomial orders for the call from* P. pipistrellus *using a window length of* 65 *samples.*

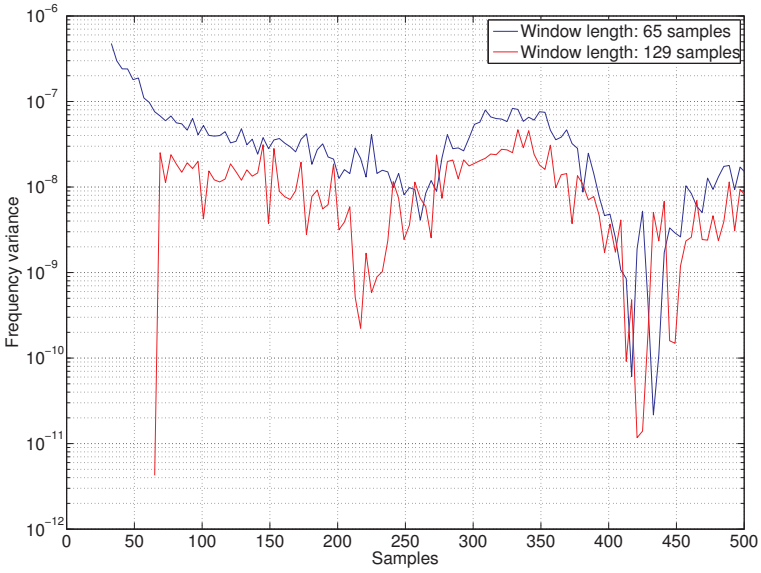


Figure 7.12: *Variance of the frequency estimate for a first-order polynomial for different window lengths when applied to the call from* P. pipistrellus.

The spectrograms show that the estimates are unable to characterise all the detail in the signal. This is most obvious in cases where the signal contains significant amounts of amplitude modulation and the constant amplitude model is unable to deal with this. However, the scale on the spectrogram shows that the harmonic components in the residual are at a level of approximately -60 dB or less, indicating that very little information has been lost.

For the purpose of comparison, the RBPF-based algorithm developed by Dubois and Davy [70] and described in Section 3.5 is applied to the call from *P. pipistrellus*. The algorithm is run using the same number of particles, *i.e.*, $10{,}000$, but only for a window length of 65 samples. For the RBPF-based algorithm, the memory requirements are proportional to the square of the window length L. This occurs due to the necessity of computing the $L \times L$ matrix $\mathbf{S}_k$ in Table 5.1 (page 107). When using a window length of 129 samples and $10{,}000$ particles, the memory requirements become prohibitive and these results are left out for this reason.

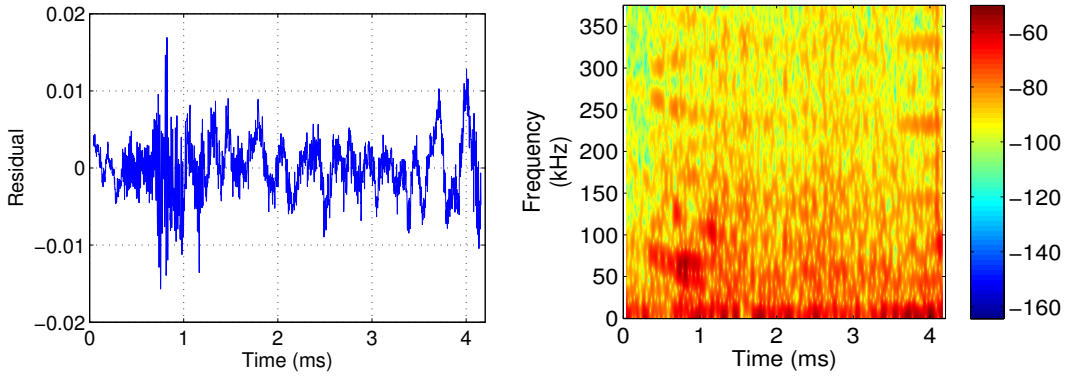
Figure 7.14 illustrates the frequency estimates obtained from Dubois and Davy's algorithm for polynomial basis orders 1 through 3. The algorithm is constrained to detecting a single tone with two harmonics resulting in two tracks begin present in the estimates. Alongside the frequency estimates, the magnitudes of the components are also displayed. It can be seen that the higher harmonic has a very low amplitude in some parts of the signal and some form of thresholding is required to test for the presence of the higher harmonic.

Similar to the residuals illustrated for the MPF-based algorithm, Figure 7.15 shows the residuals for the RBPF-based algorithm for the different polynomial orders considered here. The results are similar to those illustrated earlier with the algorithm suffering due to the constant amplitude model.

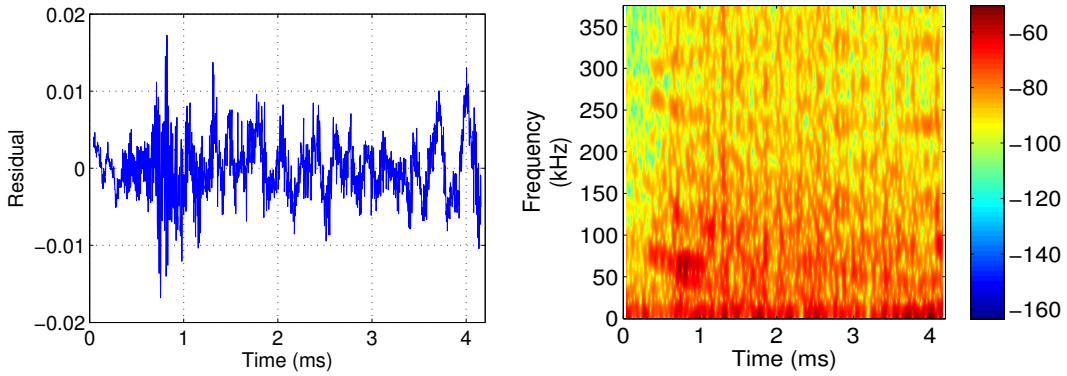
In order to compare the performance of the RBPF- and MPF-based algorithms, the sum of squared residuals is considered in Table 7.2. The table shows that the RBPF performs slightly better than the MPF. Additionally, for the RBPF, a polynomial order of 2 is seen to offer the best results, while a polynomial order of 1 is preferred in the case of the MPF. It is noted however, that the lower residuals of the RBPF are achieved only at higher computational expense.

7.2.2 Analysis of Call from *M. daubentonii*

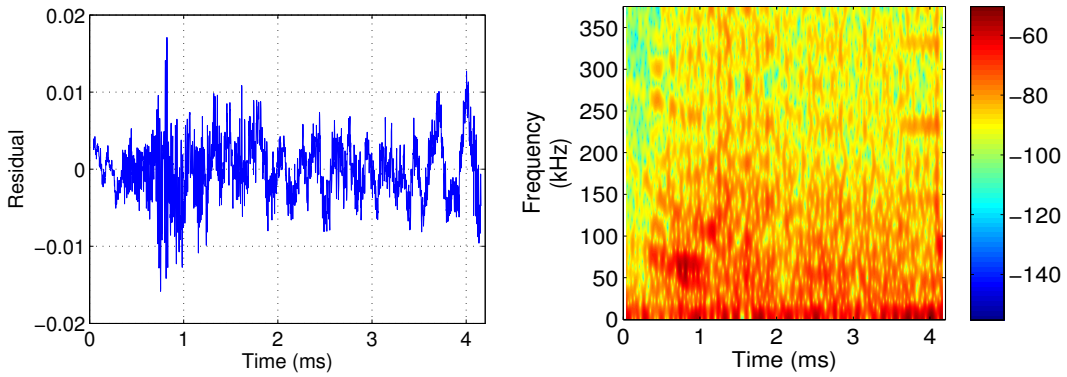
The second test call, from Daubenton's bat possesses a significant amount of amplitude modulation. Notches in the time-domain signal are seen to conincide with ambiguities in the spec-



(a)

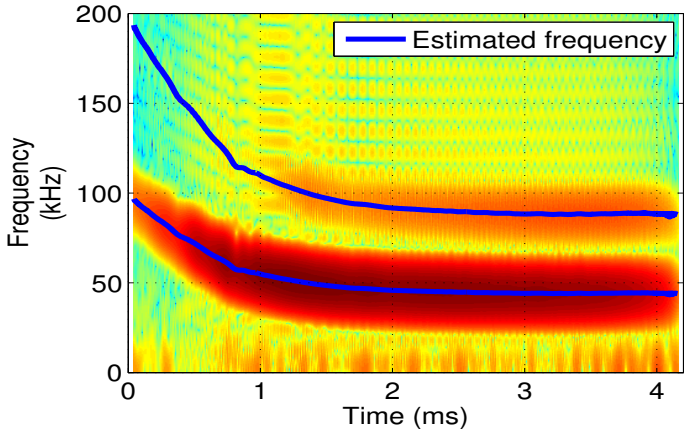


(b)

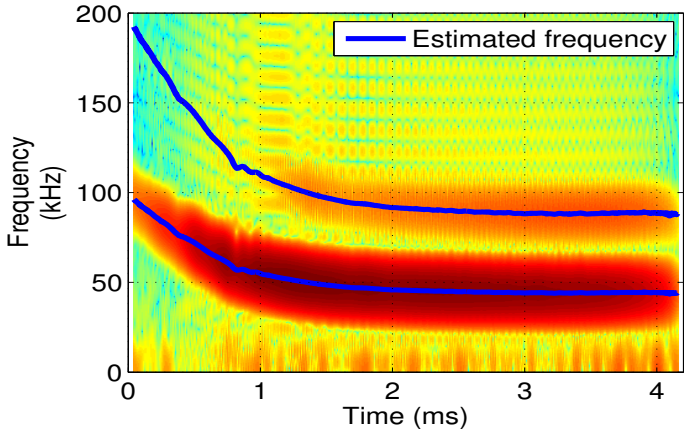


(c)

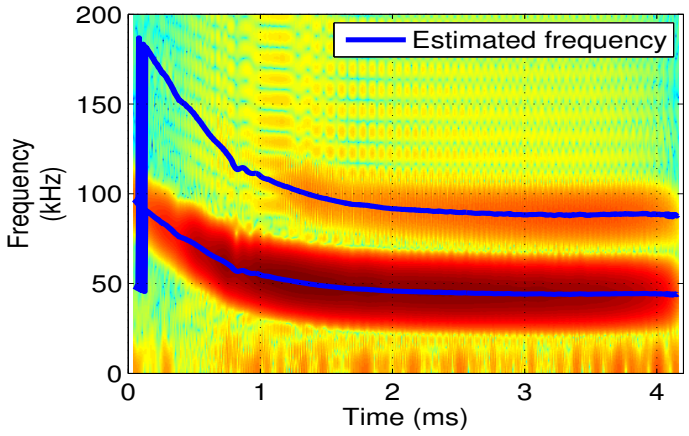
Figure 7.13: *Residuals with their spectrogram using the MPF-based algorithm for the call from* P. pipistrellus *using a window length of* 65 *samples with a polynomial basis of order* 1 *(top),* 2 *(centre) and* 3 *(bottom).*



(a)

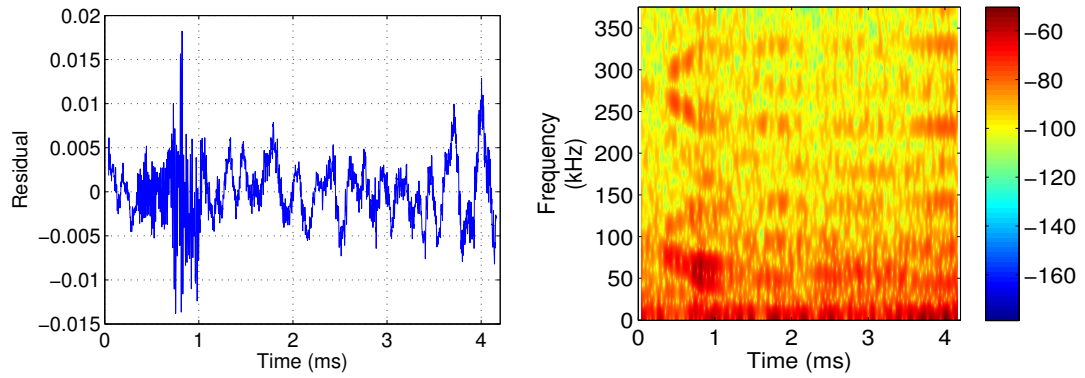


(b)

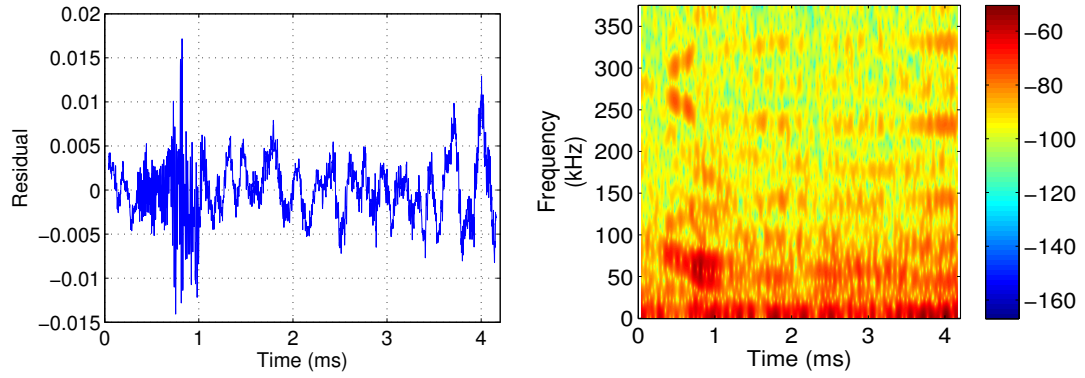


(c)

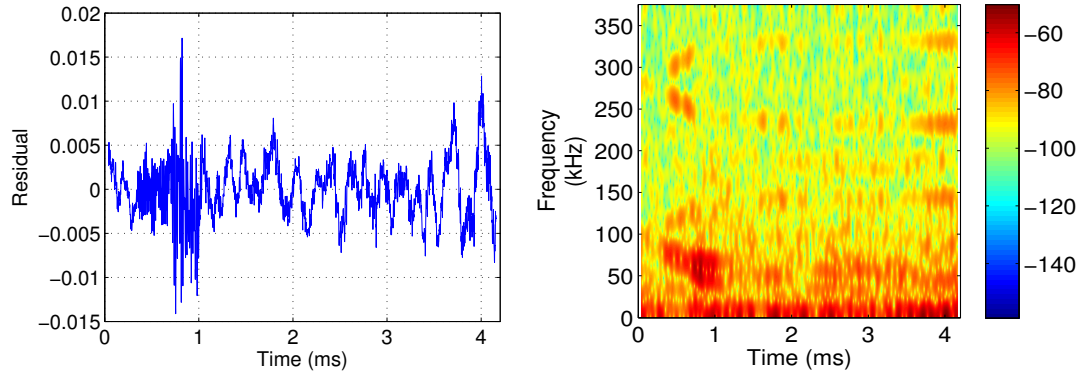
Figure 7.14: *Instantaneous frequency estimates for the call from* P. pipistrellus *using the RBPF for a window length of* 65 *samples with a polynomial basis of order* 1 *(top),* 2 *(centre) and* 3 *(bottom).*



(a)



(b)



(c)

Figure 7.15: *Residuals with their spectrogram using the RBPF-based algorithm for the call from* P. pipistrellus *using a window length of* 65 *samples with a polynomial basis of order (a)* 1*, (b)* 2 *and (c)* 3*. Spectrogram of the residuals is shown on a dB scale.*

Algorithm	Sum of squared residuals (10^{-2}) for polynomial order of:		
	1	2	3
RBPF	3.5639	3.4596	3.4619
MPF	4.0359	4.3264	4.4121

Table 7.2: *Comparison of residuals for the call from* P. pipistrellus *for the RBPF- and MPF-based algorithms using different model orders for a window length of* 65 *samples.*

trogram. Figure 7.16 shows the estimated frequency of the call using window lengths of 65 samples (left) and 129 samples (right). Considering the 65 sample window, irrespective of the polynomial order, the algorithm is unable to obtain a smooth estimate of the frequencies occurring at approximately 0.9 ms in the signal. Increasing the window length to 129 samples results in a smoother estimate of the frequencies over time and the jumps in the frequency which occur at 0.9 ms no longer occur as seen in the figure.

Examining the results from using a 129 sample window, it is seen that there is a difference between the instantaneous frequency estimates using a first-order and higher order polynomials. While the longer window reduces the variance of the frequency estimates, the first-order polynomial is unable to adequately characterise the modulation within this longer window due to the lack of model flexibility. A second-order polynomial basis, however, is capable of providing smooth estimates without losing detail.

Figure 7.17 compares the variance of the frequency for the different polynomial orders for a window length of 65 samples. As in the case of the pipistrelle call, a first-order polynomial is associated with lower variance on the frequency estimate. The figure also highlights the increase in the variance at notches in the signal, points which are associated with ambiguities in the spectrogram. Increasing the window length to 129 samples is seen to reduce the variance further as shown in Figure 7.18, indicating that the model provides a reasonable fit even within the longer window.

In addition to the variance, Figure 7.19 shows the residuals and their spectrograms for the call from *M. daubentonii*. As was seen in the residuals for the call from *P. pipistrellus*, the residuals are larger in the presence of any significant amplitude modulation. However, the spectrogram shows that the residuals still possess very little energy and appear at less than -60 dB.

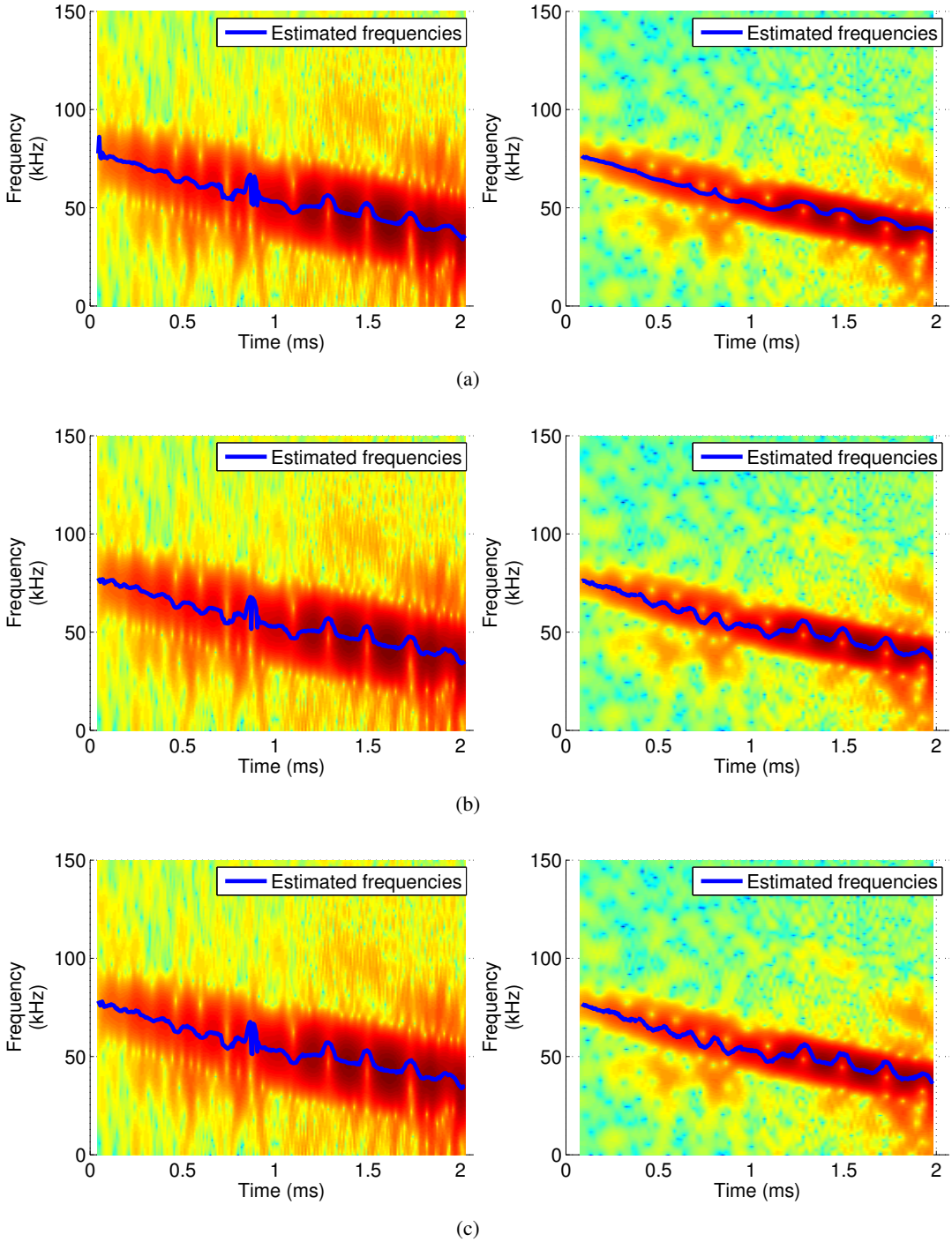


Figure 7.16: *Instantaneous frequency estimates for the call from* M. daubentonii *using the MPF using polynomial basis orders (a)* 1*, (b)* 2 *and (c)* 3*. The figures on the left demonstrate results using a window length of* 65 *samples and those on the right use a window length of* 129 *samples.*

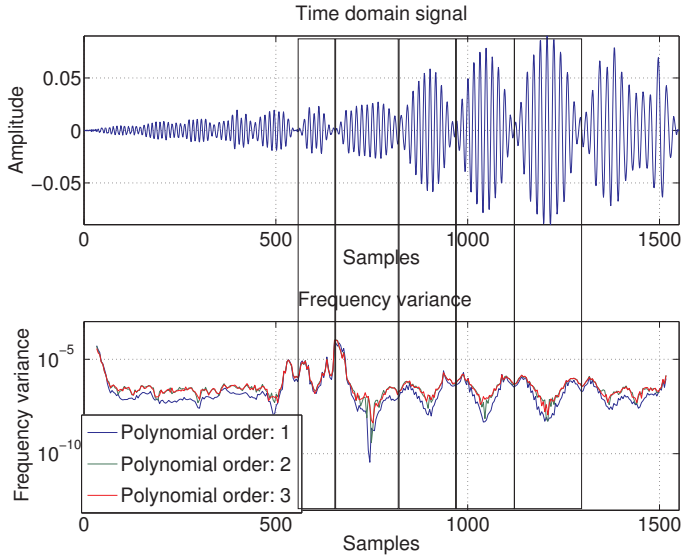


Figure 7.17: *Variance on the frequency estimate for varying polynomial orders for the call from* M. daubentonii *using a window length of* 65 *samples.*

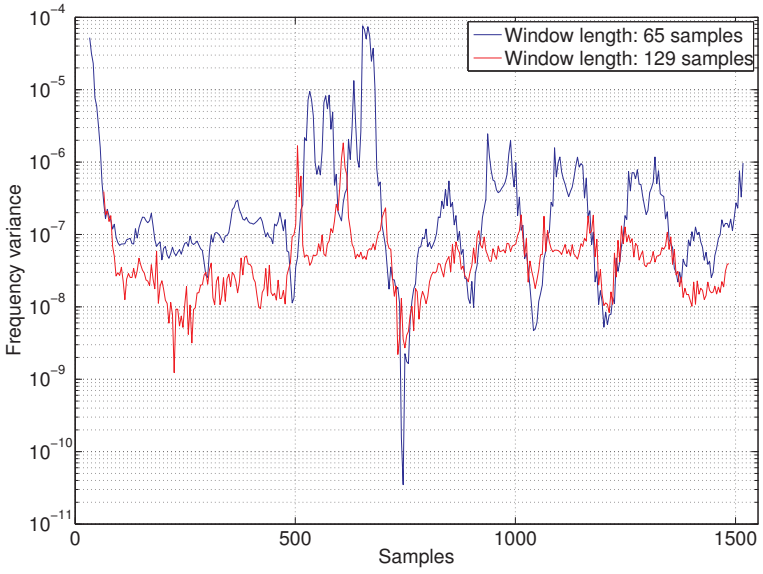
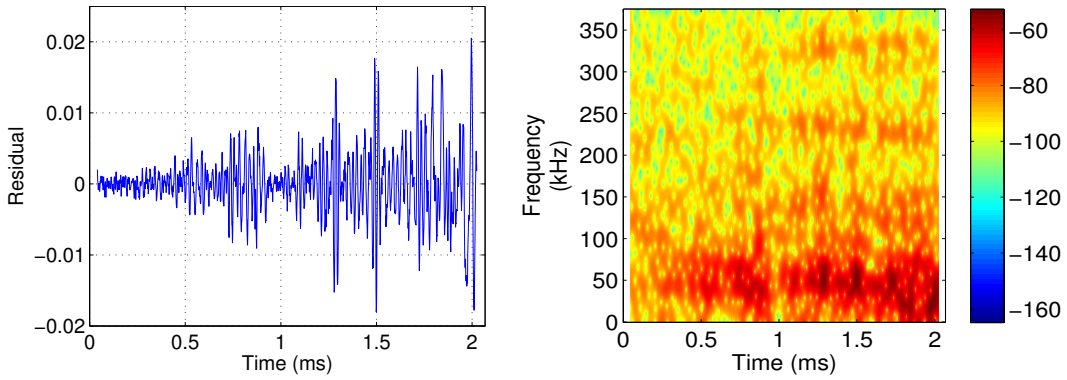
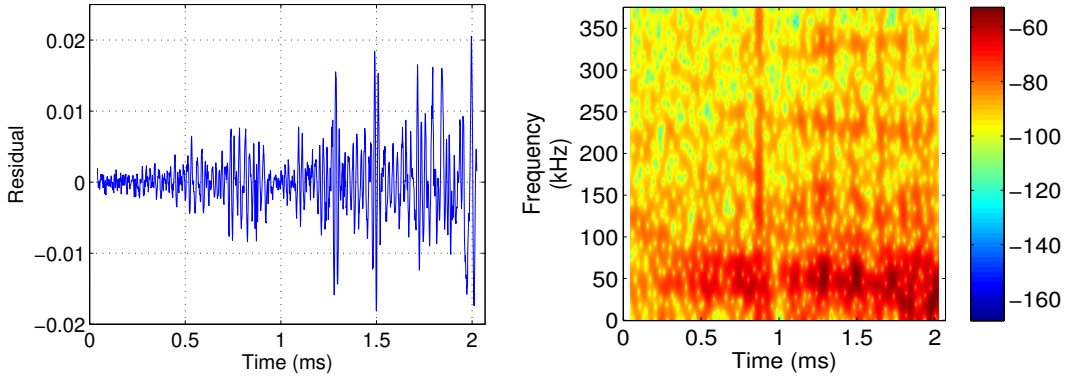


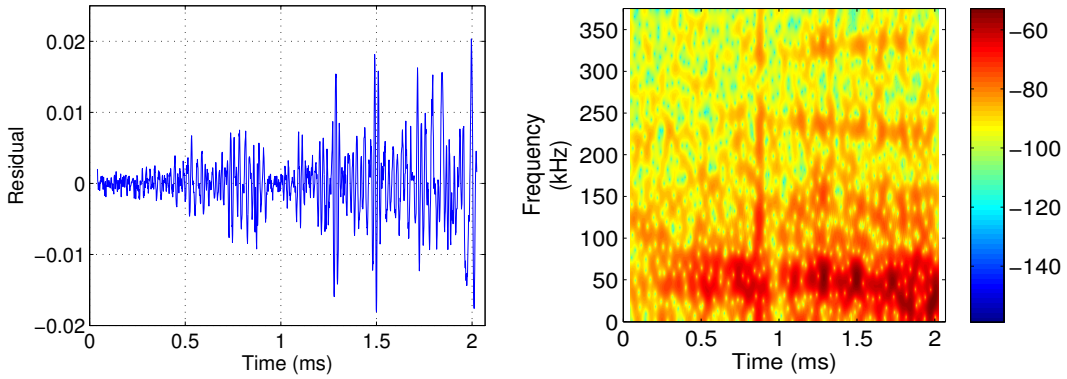
Figure 7.18: *Variance of the frequency estimate for a first-order polynomial for different window lengths when applied to the call from* M. daubentonii.



(a)



(b)



(c)

Figure 7.19: *Residuals with their spectrogram using the MPF-based algorithm for the call from* M. daubentonii *using a window length of* 65 *samples with a polynomial basis of order (a)* 1*, (b)* 2 *and (c)* 3*. Spectrogram of the residuals is shown on a dB scale.*

Algorithm	Sum of squared residuals (10^{-2}) for polynomial order of:		
	1	2	3
RBPF	3.1723	3.1743	3.2077
MPF	3.0495	3.0443	3.0734

Table 7.3: *Comparison of residuals for the call from* M. daubentonii *for the RBPF- and MPF-based algorithms using different model orders for a window length of* 65 *samples.*

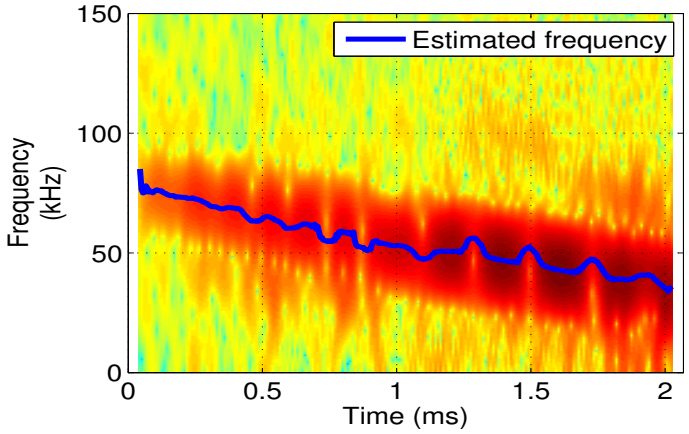
As in the analysis of the call from *P. pipistrellus*, the RBPF is applied to the call from *M. daubentonii* for the purpose of comparison. Figure 7.20 shows the results from application of the RBPF to the call using a window length of 65 samples and polynomial basis orders 1 through 3. The results from all three model orders are similar with only the linear chirp basis (polynomial order 1) showing slightly less detail compared to the other polynomial orders.

Using the RBPF, the residuals and their spectrograms for the call are shown in Figure 7.21. The residuals are not dissimilar to those of the MPF in Figure 7.19. The sum of the squared residuals for both the RBPF and MPF are presented in Table 7.3. The results here show that the residuals are slightly lower in the case of the MPF. It is also seen that in the case of the MPF, a 2^{nd} order polynomial basis is able to capture more of the detail in the signal structure.

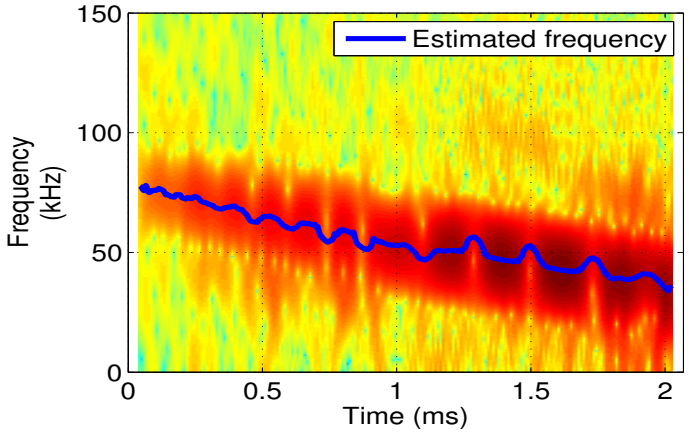
The RBPF and the MPF have been applied to two test calls to provide insight into the choice of polynomial order and window length. While a first-order polynomial basis may be adequate for most signals, the algorithms do not necessarily provide acceptable results for signals with significant frequency and amplitude modulation as is seen in the the call from Daubenton's bat in Figure 7.16(a). The choice of basis depends on the signal under consideration, with a higher order polynomial basis essential for signals possessing significant amplitude and/or frequency modulation.

7.3 Inharmonicity in Bat Echolocation Calls

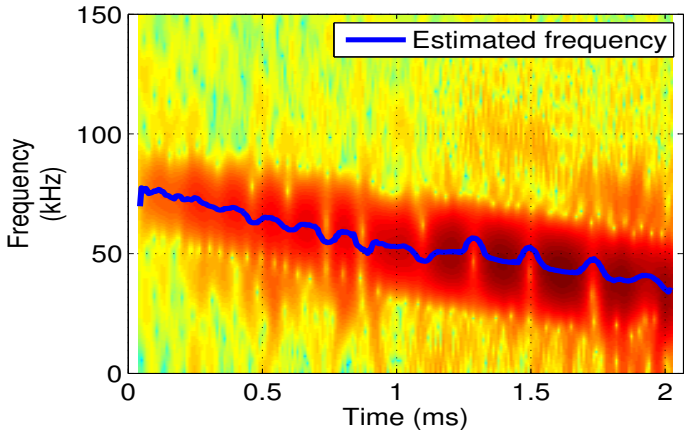
The second main issue examined in this chapter is the presence of inharmonicity in bat echolocation calls. The harmonic model described in Section 4.3.2.2 is able to account for overtones in the signal as well as deviations which may be present in those overtones. While such a model is particularly useful in modelling certain musical instruments [38], the model can also be used



(a)

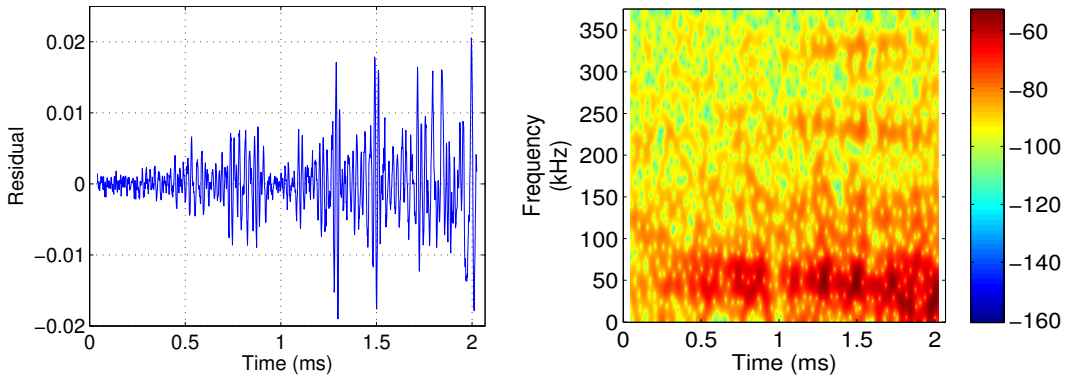


(b)

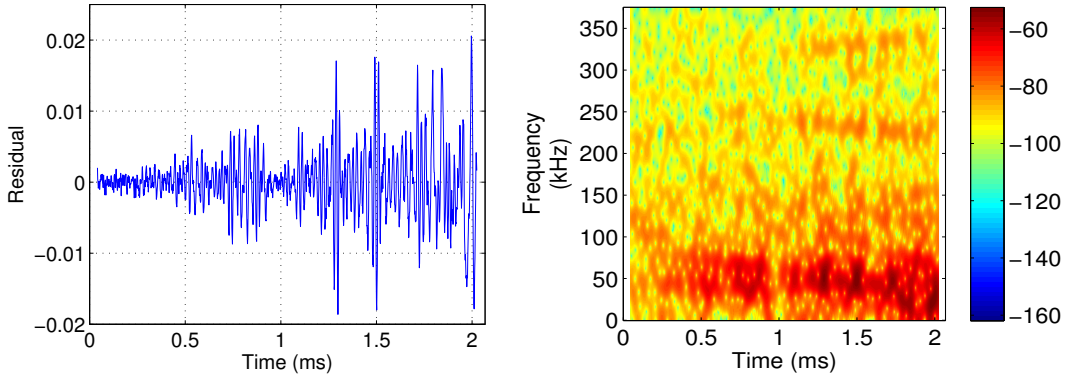


(c)

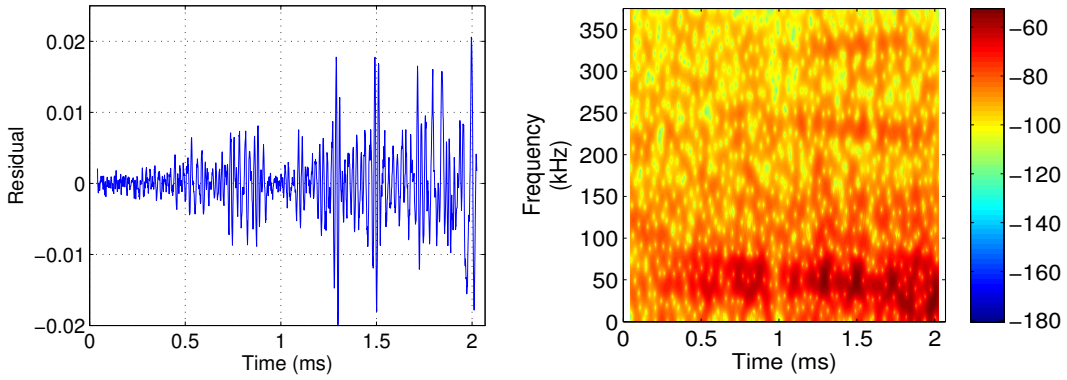
Figure 7.20: *Instantaneous frequency estimates for the call from* M. daubentonii *using the RBPF for a window length of* 65 *samples with a polynomial basis of order (a)* 1*, (b)* 2 *and (c)* 3*.*



(a)



(b)



(c)

Figure 7.21: *Residuals with their spectrogram using the RBPF-based algorithm for the call from* M. daubentonii *using a window length of* 65 *samples with a polynomial basis of order (a)* 1*, (b)* 2 *and (c)* 3*. Spectrogram of the residuals is shown on a dB scale.*

to detect deviations in the harmonics of bat echolocation calls.

Estimation of the harmonic deviation was discussed in Section 6.4, where the algorithm was set up to detect a deviation of 1%. The results presented in that section show that detection of such a deviation is possible with sufficient accuracy, even in the presence of amplitude modulation. In this section, the algorithm is applied to estimate deviations which may be present in segments of calls from six different bat species. Segments of 2000 samples (equivalent to 2.67 ms at 750 kHz sampling rate) are selected from calls such that the segment contains the fundamental and at least one overtone.

For each call considered, the overtone deviation is estimated using 100 independent runs. The results are then presented as the mean and the standard deviation averaged over all runs. Figure 7.22 shows the results from a single call from each of six selected species: *Pipistrellus capensis* (Smith, 1829), *Pipistrellus Kuhlii* (Kuhl, 1817), *P. pipistrellus*, *Pipistrellus pygmaeus* (Leach, 1825), *Rhinolophus capensis* (Lichtenstein, 1823) and *Tadarida Aegyptiacus* (Geoffroy, 1818). Although the number of species considered here is small, different types of call structures are covered between these six species, *i.e.*, both FM and CF calls are studied.

In some of the calls, the overtone deviation appears to indicate a marginal bias. However, the bias is considerably smaller than the 1% estimation accuracy established for the method. These results go some way to showing that there is no significant, detectable deviation in the calls.

To ensure that this lack of overtone deviation is consistent, three calls are selected from each species and tested. Table 7.4 summarises the average mean and standard deviation from analysis of the set of call segments from each species. The call segments are taken from calls recorded from different individuals, thus ensuring sufficient diversity. From the table, it is clear that these calls do not contain any measurable overtone deviation. It can thus be assumed that overtone deviation does not play any significant role in the structure of the calls considered here, and by extension, the presence of overtone deviation in calls from other species would be unexpected.

7.4 Analysis of Feeding Buzz Sequences

This section addresses application of the frequency estimation algorithm to the study of bat echolocation calls constituting the *feeding buzz*. These calls are seen just prior to prey capture and differ from the other calls in some respects. The context of the feeding buzz is discussed

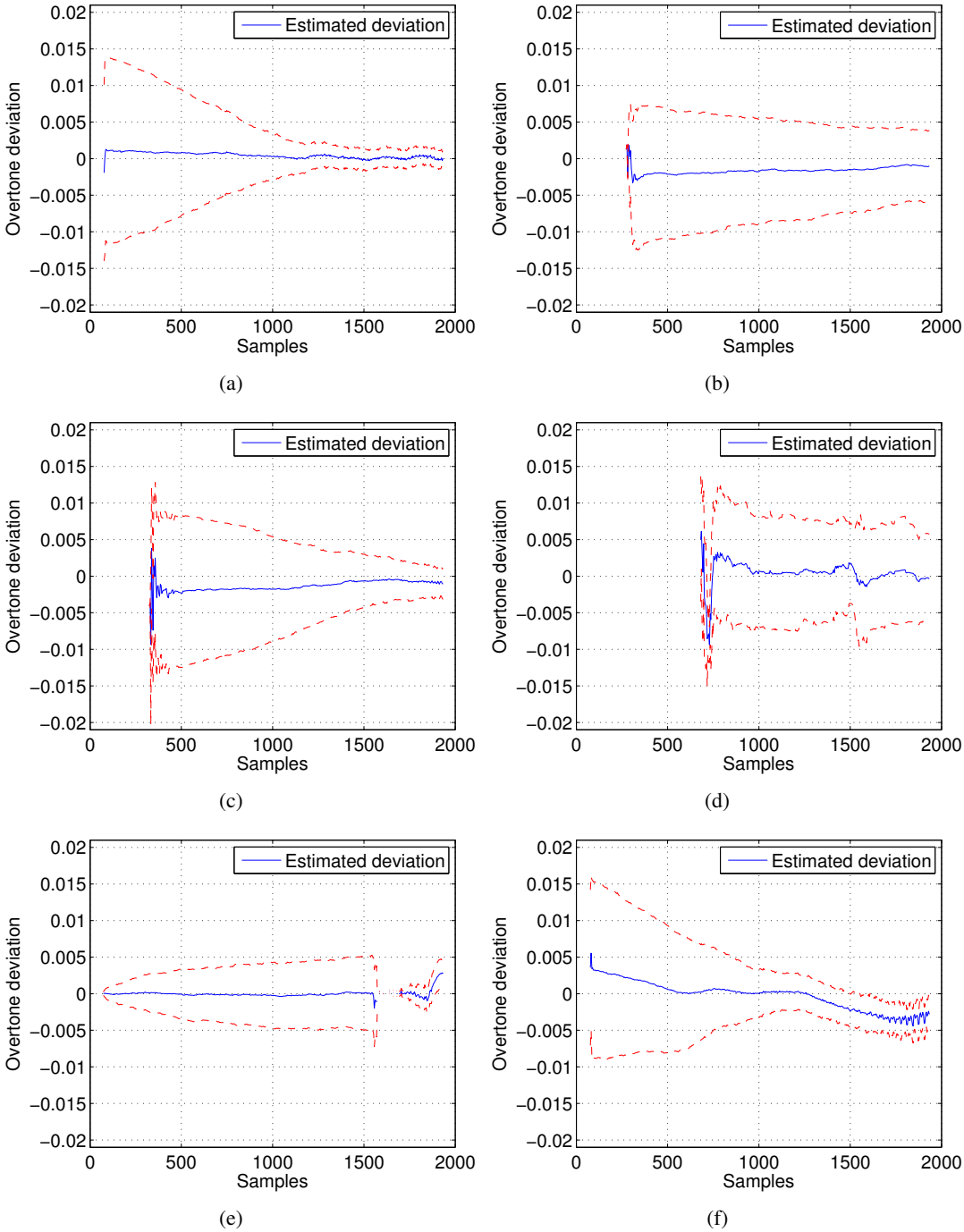


Figure 7.22: *Mean and standard deviation of overtone deviation parameter for* (a) P. capensis (b) P. Kuhlii (c) P. pipistrellus (d) P. pygmaeus (e) R. capensis *and* (f) T. Aegyptiacus.

Species	Mean (10^{-3}) ± Standard deviation (10^{-3})		
	Call 1	Call 2	Call 3
P. capensis	$+0.449 \pm 4.850$	-1.060 ± 6.069	-1.209 ± 3.319
P. Kuhlii	-1.616 ± 6.915	-1.237 ± 6.225	-0.440 ± 8.549
P. pipistrellus	-1.359 ± 6.191	-0.652 ± 6.712	-2.386 ± 8.706
P. pygmaeus	$+0.103 \pm 6.302$	$+4.009 \pm 4.286$	-0.252 ± 1.362
R. capensis	$+0.108 \pm 3.799$	-0.180 ± 3.902	-0.009 ± 3.336
T. Aegyptiacus	-0.271 ± 5.139	$+1.077 \pm 3.046$	-2.811 ± 5.118

Table 7.4: *Estimated mean and standard deviation of overtone deviation estimates for different bat species.*

first, and subsequently, the frequency estimation algorithm is applied to the calls. Due to the structure of the feeding buzz calls, some amount of hyper-parameter tuning is required to obtain reasonable results and this will be discussed as well.

7.4.1 Overview of Feeding Buzz Calls

When bats are hunting for prey, the structure of the echolocation call is tuned to the task the bat is focussing on. As discussed earlier in Section 2.4, the bat may be actively looking for prey in the *search* phase, it may have located prey and started flying towards the target in the *approach* phase, or it may be in the final stages of capturing the target in the *terminal* phase, with these different phases illustrated in Figure 2.1 (see page 30). The requirements of each stage are different, and consequently, the structure of the calls is adapted to reflect the requirements.

In the search phase, the calls are typically of longer duration, on the order of several milliseconds, and usually have lower bandwidth compared to calls from the other stages. This may arise from the fact that lower frequency calls are attenuated to a lesser degree, allowing the bat to detect targets which are further away [107]. Additionally, the bat does not require highly accurate directional information about the target, but merely needs to detect it during this stage. It has also been suggested that the longer duration of the call allows the bat to detect modulation glints arising from flapping wings of the target, thus allowing the bat to possibly obtain further information about the target [63].

Once the bat has located its prey, it starts tracking the target while flying towards it. In this

case, directional cues become more important, and the duration of the call reduces with a simultaneous increase in the bandwidth. Reducing the duration of the calls decreases the chance of pulse-echo overlap which is important if the bat is feeding in the vicinity of vegetation.

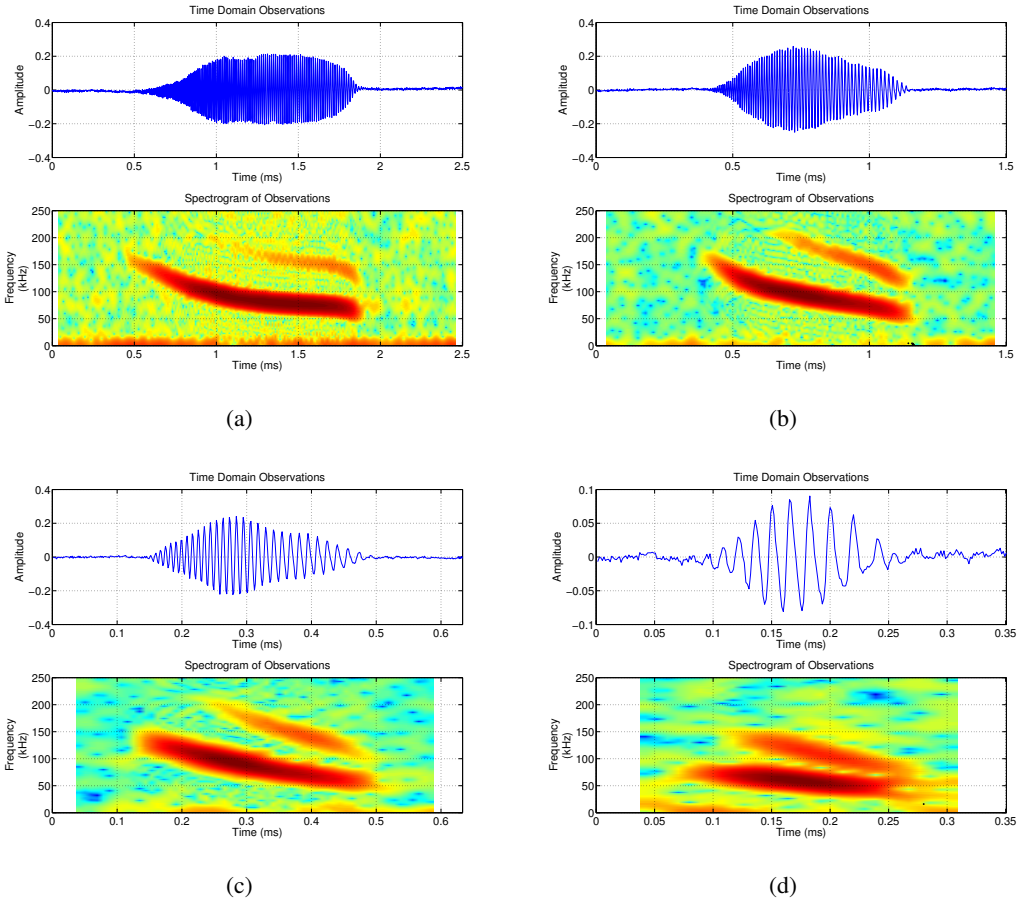


Figure 7.23: *Calls from the feeding buzz of a pipistrelle with the STFT used to obtain the time-frequency representation.*

In the final stage of prey pursuit, the bat adapts the calls to provide a high degree of localisation information. This is achieved by reducing the call duration even further, increasing the bandwidth of the call, and increasing the pulse repetition frequency (PRF) resulting in the feeding buzz. The feeding buzz gets its name from the fact that the calls in this segment are extremely close together, such that they sound like a buzz when played at audible frequencies.

The duration of these calls can be as short as 0.2 ms (150 samples at 750 kHz) making the analysis of these calls extremely difficult. The traditional use of the STFT by biologists for examination of the feeding buzz does not permit an accurate examination of these signals due to

the presence of both frequency and amplitude modulation within the extremely short duration of the calls. Figure 7.23 shows recordings of calls from the feeding buzz of a pipistrelle along with the spectrograms. While some detail is discernible in the longer calls, very poor characterisation is achieved for the very short duration calls.

7.4.2 Deviations from the Signal Model and Hyper-parameter Tuning

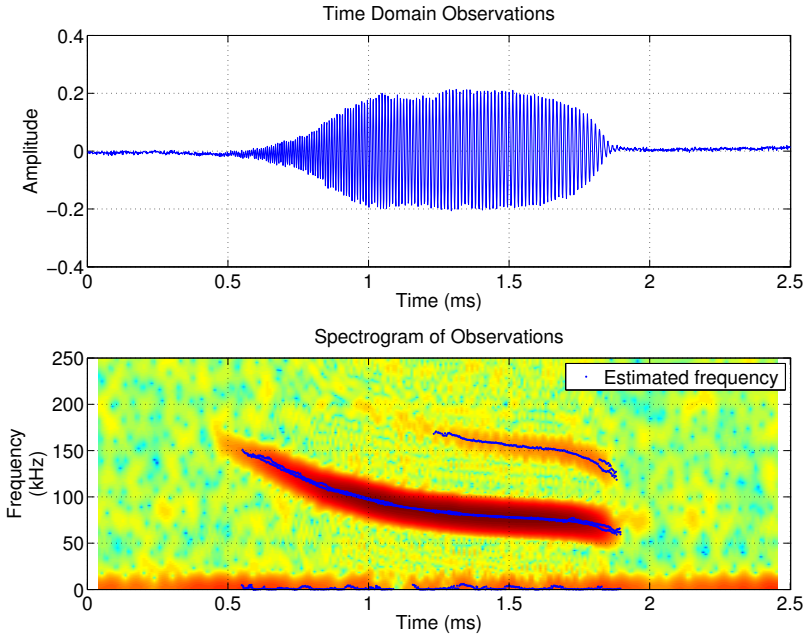
Figure 7.24 shows the results from application of the frequency estimation algorithm applied to each of the calls. In these examples, the filter is run using 1000 particles and a window length of 65 samples, corresponding to 0.087 ms. By contrast, the shortest echolocation call considered here has a duration of 0.2 ms. In this context, a window length of 65 samples may be considered long, however, reducing the window length will increase the variance on the frequency estimates which is undesirable.

The figure shows that, rather than a single frequency, the algorithm detects two or more closely spaced tones in place of the fundamental tone. The detection of multiple tones arises from the violation of the assumptions made in the signal model. In particular, the signal model assumes a constant amplitude of the tone within the sliding window; the algorithm attempts to account for the deviations in the observed signal, *i.e.*, the amplitude modulation, by adding an extra tone which gives rise to amplitude modulation as a result of interference between the tones.

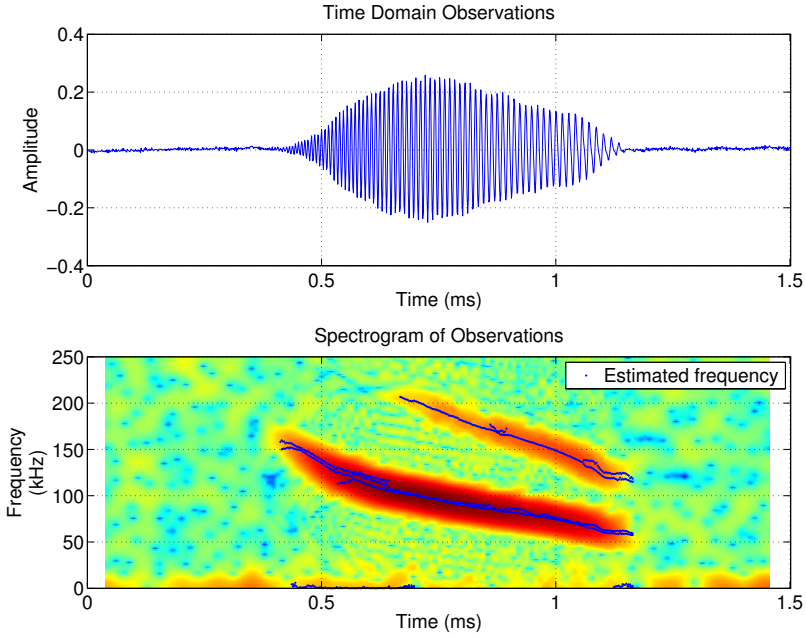
In addition to the problem described above, the algorithm may occasionally suffer from spurious detection of tones. These spurious detections arise from the fact that the observed signal does not strictly conform to the assumptions made in the signal model. For example, the observation noise spectrum, in part shaped by the transducer characteristics of the recording device, may be coloured, rather than white. Additionally, when recording bat echolocation calls in the field, weak reflections of calls from surrounding vegetation, not corresponding to the signal of interest, may also be picked up in addition to various other environmental sounds. These weak components may be detected by the algorithm and appear as spurious detections.

The approach adopted here to address these issues is to tune the hyper-parameters of the algorithm, which involves incorporating additional constraints in the prior (and sampling) distributions.

When examining the bandwidth of echolocation calls, it is seen that calls do not occur at less than 20 kHz; by contrast, a great deal of environmental noise occurs below this frequency. Thus,

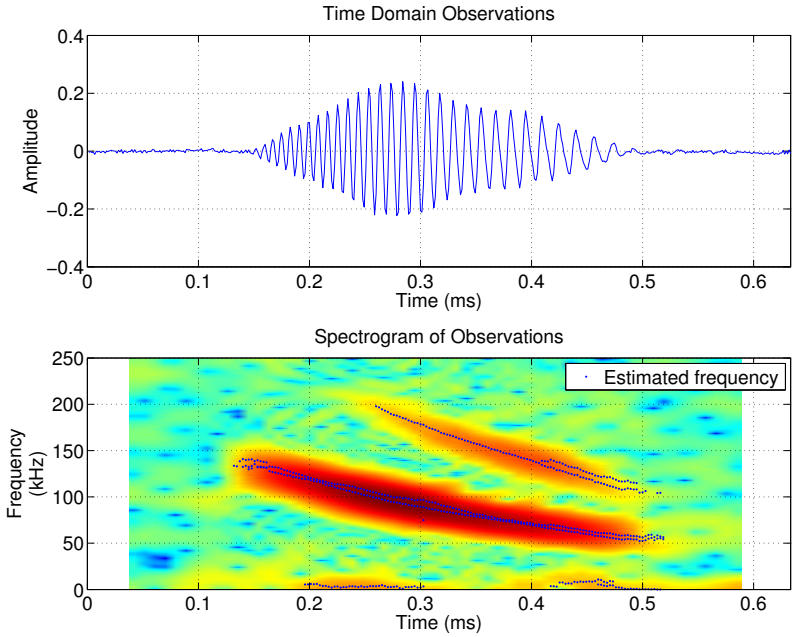


(a)

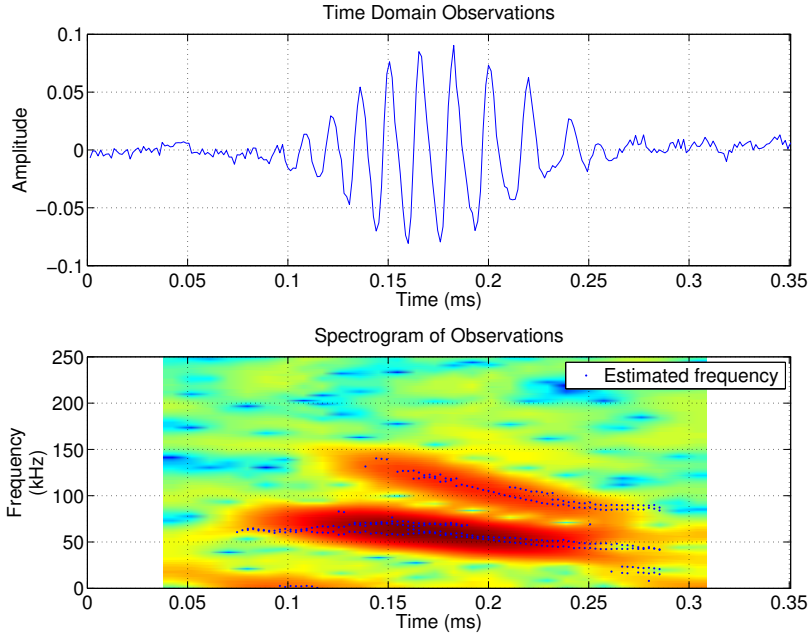


(b)

Figure 7.24: *Application of the frequency estimation algorithm to the feeding buzz calls shows that the algorithm has difficulty estimating single tones in the presence of amplitude modulation.*



(c)



(d)

Figure 7.24: *Application of the frequency estimation algorithm to the feeding buzz calls shows that the algorithm has difficulty estimating single tones in the presence of amplitude modulation.*

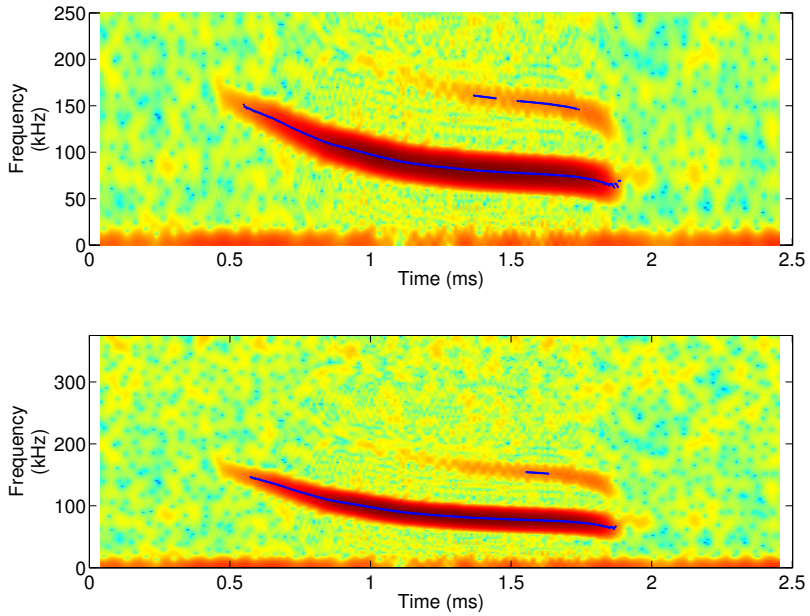
specifying $f_{min} = 20$ kHz in the prior distribution allows the algorithm to ignore a significant source of undesirable noise.

To prevent the algorithm from detecting a single tone as multiple closely spaced components, an additional constraint is imposed on the sampling and prior distributions such that a minimum frequency spacing, f_{null}, is required between any two tones. This introduction of nulls in the distribution prevents the algorithm from proposing a new tone with a frequency close to that of an existing component. As a result, the algorithm is forced to estimate the call as a single tone. The obvious disadvantage of this is the inability of the algorithm to detect closely separated components, however, it provides an easy means for forcing the algorithm to detect only a single component.

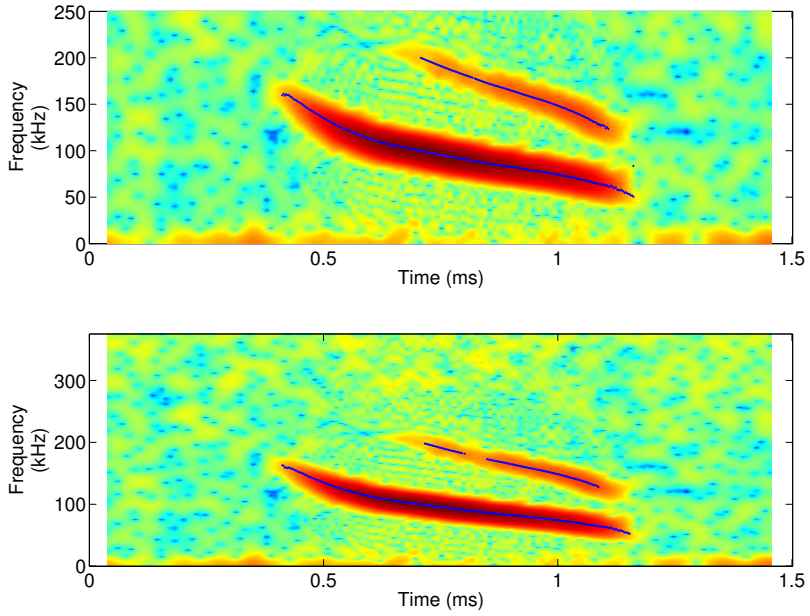
In order to reduce the number of spurious detections, it is necessary to tune the algorithm to ignore very weak harmonic components. This is done in practice by modifying the distribution on δ_k^2 to a lower-truncated inverse gamma distribution. Since δ_k^2 is representative of the SNR of the signal (see Section 4.4.2), a lower limit, ${\delta_k}^2_{min}$, provides a means of specifying a minimum desired SNR of the estimates, such that components which do not meet the SNR requirement will be ignored. An examination of the likelihood function in equation (4.40) (on page 89) shows that the term $\delta_k^{2O_k}$ acts as a penalty term. Any additional tone must offer sufficient benefit to overcome the penalty generated by its introduction. Enforcing a minimum value then forces that new tone to exceed a certain minimum SNR, thus reducing the instances of spurious detections.

The algorithm is rerun using these modified prior and sampling distributions for the parameters. The shorter duration calls possess a large bandwidth within the sliding window. Consequently, f_{null} needs to be sufficiently large in order to suppress multiple closely spaced tones, and a value of $f_{null} = 22.5$ kHz is used. Two sets of tests are run with ${\delta_k}^2_{min} = 10^2$ and ${\delta_k}^2_{min} = 5 \times 10^4$ to illustrate the effect of this threshold. The prior on δ_k^2 is uninformative and the hyper-parameters $(\alpha_\delta, \beta_\delta)$ can take on a large range of values. As discussed in Section 4.4.2, $\beta_\delta = 2$ specifies a distribution with infinite variance, and α_δ is chosen an order of magnitude larger than ${\delta_k}^2_{min}$.

The results from the algorithm after tuning the parameters of the priors are shown in Figure 7.25. In each case, the upper plot shows the results from specifying ${\delta_k}^2_{min} = 100$ and the lower, ${\delta_k}^2_{min} = 5 \times 10^4$. The figure shows that the tones are now estimated as single com-

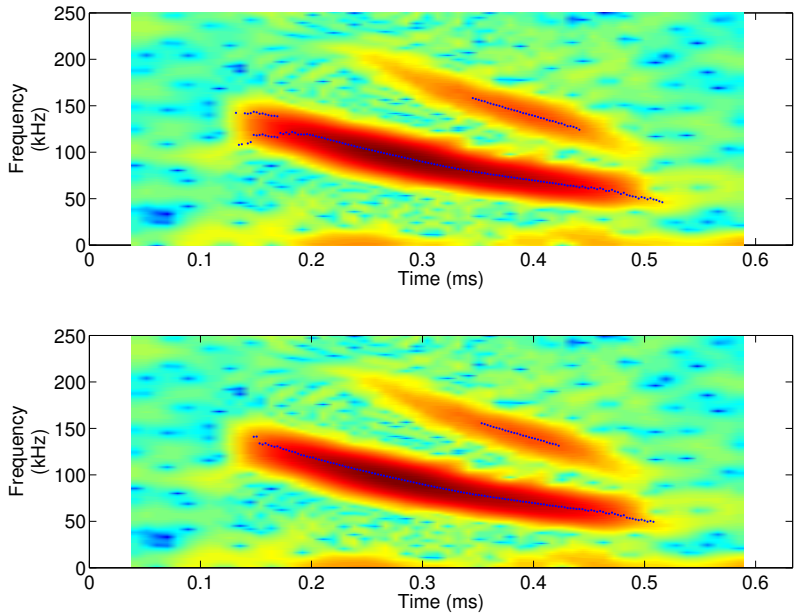


(a)

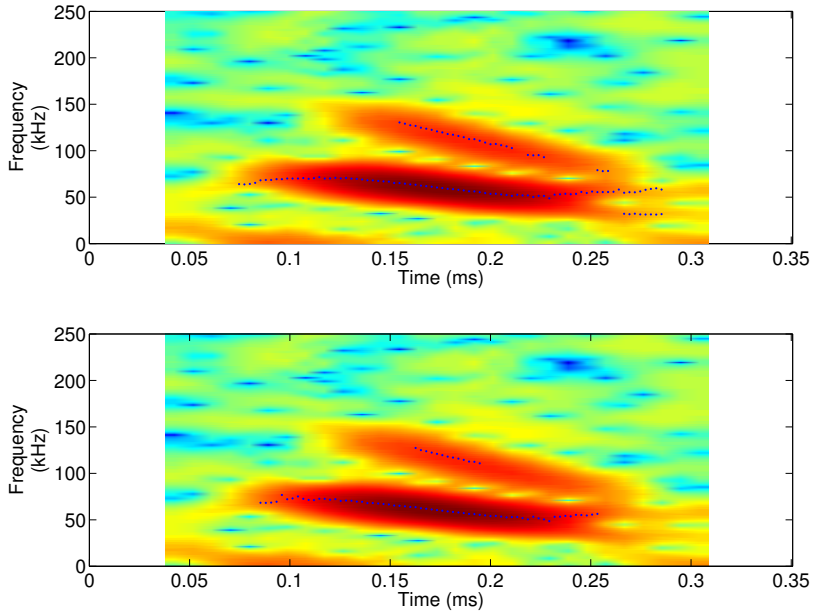


(b)

Figure 7.25: *Application of the frequency estimation algorithm using the modified distributions leads to improved results.*



(c)



(d)

Figure 7.25: *Application of the frequency estimation algorithm using the modified distributions leads to improved results.*

ponents and a large number of the spurious detections which appear in Figure 7.24 are also successfully removed.

While thresholding the value of δ_k^2 offers some benefits, it can also result in missed detections if ${\delta_{k}}_{min}^2$ is not carefully chosen. This is best illustrated in Figures 7.25(a) and 7.25(d) where portions of the overtone are not detected. An alternative method for detecting spurious frequency components is to convert the frequency estimates in the time-frequency representation into a set of frequency tracks using data association methods [50, 108], and subsequently removing tracks which exist only over a brief period of time.

The results show that the frequency estimation algorithm is capable of providing a suitable time-frequency representation of extremely short duration signals as seen in the case of echolocation calls from the feeding buzz. The method can thus provide biologists an extra tool in the analysis of calls.

7.5 Summary

Analysis of bat echolocation calls has traditionally relied on the STFT. The purpose of developing a sequential frequency estimator has been to provide a tool which is capable of improving upon the time-frequency representation of the STFT and similar methods by taking signal structure into account.

This chapter has illustrated how the sequential frequency estimator may be applied to echolocation calls successfully. Sections 7.1 and 7.2 examined the choice of the polynomial basis function for a synthetic test signal and two echolocation calls. It is seen that the choice of polynomial order depends on the signal under consideration, although no apparent improvement is obtained beyond a second-order polynomial for the window lengths considered here.

Section 7.3 discussed the presence of inharmonicity in echolocation calls in several different bat species. The results indicate that the calls do not contain any significant overtone deviation as was shown from the analysis of multiple calls. These results indicate that it can be reasonably assumed that overtone deviation is not present in bat calls in general.

The concluding section of this chapter considered the analysis of calls from the feeding buzz portion of a call sequence. The feeding buzz consists of extremely short duration AM-FM calls occurring just prior to prey capture. These calls can be as short as 0.2 ms and as a result,

STFT based analysis is unable to provide an acceptable time-frequency representation of the call. The sequential frequency algorithm is also seen to face problems when applied to calls of such short duration, however, it is possible to tune the priors which yields promising results as shown. Such hyper-parameter tuning may be necessary when analysing real signals which may not strictly adhere to the assumptions in the signal model.

Chapter 8
Conclusions

This chapter aims to summarise the outcomes of the research detailed in this thesis. The following section provides an overview of the work carried out herein. This is followed by a discussion of the achievements and contributions of the research. The concluding section looks at some of the limitations of the algorithms and methods developed in the previous chapters and also examines directions for future research.

8.1 Summary

The driving aim behind the work carried out here has been the development of a novel algorithm for improved time-frequency analysis of bat echolocation signals. A parametric approach is adopted here where short segments of a call are fit to a time-varying general harmonic model (GHM) which is used to track the frequencies of a call as they evolve over time.

To allow for a high degree of flexibility, the harmonic model allows the signal to be decomposed using an n^{th} order polynomial chirp basis, rather than merely unmodulated tones. The use of the polynomial basis allows for characterisation of greater detail within each windowed segment of the signal. While the GHM is used to model the signal within a sliding window, state space equations are used to model the sequential evolution of the parameters. A particle filters is chosen to implement the resulting recursive Bayesian filter.

When dealing with a linear and Gaussian class of problems, it is possible to marginalise the parameters from the posterior distribution. The Rao-Blackwellised particle filter (RBPF) offers a framework to perform parameter marginalisation. A new method of marginalisation is developed in this thesis giving rise to what is referred to as the marginalised particle filter (MPF). A comparison between the RBPF and the MPF has been considered here. Using the MPF framework, a sequential frequency estimation algorithm is developed. This new algorithm is compared with an existing frequency estimation algorithm which is based on the RBPF.

When dealing with a time-varying number of frequencies in a signal, it is possible to track the

number of fundamental tones and harmonics using either a jump Markov system (JMS) or a reversible jump MCMC (RJMCMC) sampler. While the RJMCMC sampler is more computationally expensive, it has certain advantages over the JMS. The construction of efficient reversible jump schemes is considered here to reduce the computational burden of the algorithm without adversely affecting the quality of the frequency estimates obtained from the filter.

The penultimate chapter of this thesis discusses the application of the MPF-based frequency estimation algorithm to the analysis of bat echolocation calls. Calls from different species are analysed, and the topic of hyper-parameter tuning is introduced to improve the quality of the state estimates. The analysis examines the presence of inharmonicity in echolocation calls from several different species, and also considers analysis of extremely short duration feeding buzz calls.

8.2 Contributions of the Work

This section summarises the main contributions of the work carried out in this thesis.

- **Development of a marginalised particle filter:**

  A new marginalised form of the particle filter has been derived in this thesis, termed the marginalised particle filter (MPF). The MPF allows for marginalisation by using an independence assumption which results in a tractable integral of the posterior distribution. The MPF can be used for a wide variety of problems which fit the general linear model (GLM) framework and is specifically used here to implement a sequential frequency estimation algorithm.

  A comparison is carried out between two marginalised forms of the particle filter: the MPF and the Rao-Blackwellised particle filter (RBPF). The RBPF is a well established marginalised form of the particle filter which has seen widespread application in a wide variety of problems. A comparison between the computational complexity shows that the RBPF has a significantly higher complexity as compared to the MPF. It is also established that the estimates produced by the MPF exhibit only a slightly higher mean squared error (MSE) despite the MPF being significantly less computationally demanding.

- **Sequential frequency estimation algorithm:**

  A new sequential frequency estimation algorithm is developed here using the MPF framework. RJMCMC samplers have been previously used to estimate a time-varying number

of tones and harmonics in a batch offline scenario. The algorithm developed here builds upon these methods by extending their application to a sequential framework.

An algorithm developed by Dubois and Davy [70] has been used to track a time-varying number of harmonics in a signal using the RBPF. The authors develop a rigorous model and framework and apply the algorithm successfully to analyse real music signals. The performance of this algorithm is compared with the MPF-based sequential frequency estimator. In the comparison, it is shown that the new algorithm derived in this work compares favourably with the Dubois and Davy's algorithm. A comparison between the algorithms shows that the RJMCMC sampler offers significant benefits in comparison to a JMS which is used by Dubois and Davy.

- **Efficient RJMCMC moves:**

  The reversible jump sampler used in the sequential frequency estimator, termed the two-jump move (TJM), is based on an algorithm developed by Davy *et al*. This algorithm uses multiple reversible jump moves to determine the number of tones and harmonics in the signal leading to high computational complexity. Two alternative reversible jump schemes are developed here. These jump schemes take advantage of the smoothly varying structure of bat echolocation calls to reduce the complexity of the algorithm. The resulting modified two-jump move (MTJM) and combined jump move (CJM) algorithms are compared with the TJM and shown to perform just as well, but without the additional burden associated with the TJM.

- **Analysis of bat echolocation calls:**

  The work developed in this thesis concludes with an analysis of synthetic and real AM-FM signals. It is shown that a linear chirp basis performs significantly better than a purely sinusoidal or a higher order chirp basis when dealing with FM signals. Since it is only slightly more complex than the sinusoidal basis, it is shown to be a good candidate basis for the analysis of FM signals.

  In the analysis of bat echolocation calls, estimation of inharmonicity is considered and tested in calls from different species. The results indicate that echolocation calls do not contain any significant measurable inharmonicity. While the species examined here form a small subset of the total number, it can be reasonably assumed that inharmonicity does not play a significant role in the structure of echolocation calls.

  The analysis also includes calls which contain significant amounts of amplitude as well

as frequency modulation. Under such circumstances, the signal violates the assumptions in the signal model leading to erroneous frequency estimates. The quality of the estimates can be improved through tuning of the algorithm which requires imposing additional constraints. Using these additional constraints, the algorithm is applied successfully to the analysis of extremely short duration signals from the feeding buzz portion of a sequence of bat calls.

8.3 Limitations and Future Work

This final section of the thesis examines the main limitations of the work presented here and considers the scope for further research.

- **Limitations of the model:**

  Due to the adoption of a harmonic basis, the method cannot be used meaningfully to analyse the spectrum of impulse-type signals. Thus, while the method may also be useful for analysing, for example, music signals, it may perform poorly when applied to speech analysis due to the presence of broadband components arising from, for example, fricatives.

  One of the limitations of the general harmonic model (GHM) used here is its inability to account for amplitude modulation within the window. As a result, in the presence of any significant amplitude modulation, the algorithm may detect multiple closely spaced tones rather than a single amplitude modulated tone. This is seen in the analysis of echolocation calls in Chapter 7 where it becomes necessary to impose further constraints on the prior distributions to ensure that only a single tone is detected. Additionally, the algorithm developed here implicitly assumes that the data is windowed with a rectangular window. Rectangular windows have poor side-lobe suppression and better results may be obtained from the use of other windows.

- **Limitations of the MPF:**

  The derivation of the MPF in Chapter 5 relies on an independence assumption to evaluate the integral. While the impact of the assumption is small in the application of the MPF to frequency estimation, the same may not necessarily be true in other applications. Additionally, if the observation noise variance is known, then the RBPF offers a better alternative since it performs the exact integral, as opposed to the approximation that the

MPF relies on.

- **Constraints of the reversible jump schemes:**

  The TJM proposed in the literature for estimation of an unknown number of tones and harmonics turns out to be computationally expensive. In constructing the alternative MTJM and CJM for detecting the number of tones and harmonics, an assumption of a smoothly varying signal is made. An examination of bat echolocation calls indicates that this assumption is not unreasonable here. However, in analysis of other signals, the use of the MTJM or the CJM may result in an estimator that is slower to converge as was discussed in Chapter 6.

- **Extraction of echolocation calls:**

  One aspect of future work which is only touched upon here is *data association* – the process of converting the instantaneous frequency estimates into a set of frequency tracks. This is considered, for example, in the work by Clark *et al*. discussed in Chapter 2. Several data association techniques exist for the conversion of the instantaneous frequency estimates into a set of tracks. This conversion can be aided by the use of a polynomial chirp basis rather than a sinusoidal basis. In the case of the polynomial basis, it is possible to perform discrimination by considering the higher order terms, thus allowing improved separation of overlapping frequency components.

- **Further analysis of echolocation calls:**

  The frequency estimation algorithm developed here provides a means for performing further analysis of echolocation calls. Since the algorithm is capable of estimating the precise frequencies of the call, it is possible to analyse call sequences for trends in the call parameters within an individual, as well as within and across species. Such a study would serve to quantify the changes in call structure and move away from a merely descriptive analysis of the calls.

- **Application to other signals:**

  The application of the method has been restricted to the analysis of echolocation calls in this work. Its extension to the analysis of music and speech signals may require modifications to the signal model, for example, using damped sinusoids, or incorporation of model switching behaviour to deal with fricative sounds. These modifications, however, would increase the range of signals which can be analysed with the method.

Appendix A
Alternative Derivation of Marginalised Particle Filter for Conditionally Linear Gaussian State Space Models

This section examines an alternative approach to the one considered in Sections 4.4.4 and 5.1 in developing a marginalised form of the particle filter. As in the case of the MPF derived earlier in this thesis, this alternative approach still applies to conditionally linear and Gaussian state spaces.

Consider the state space equations:

$$\boldsymbol{\theta}_k = f(\boldsymbol{\theta}_{k-1}) + \mathbf{v}_{(\theta),k-1} \tag{A.1}$$

$$\mathbf{x}_k = \mathbf{G}_k \mathbf{a}_k + \mathbf{e}_k \tag{A.2}$$

where $\boldsymbol{\theta}_k$ represents the state at time step k which is to be estimated, $f(\cdot)$ denotes a function for the state update and $\mathbf{v}_{(\theta),k-1}$ is the vector of process noise; $\mathbf{x}_k$ is the observation vector, $\mathbf{G}_k = g(\boldsymbol{\theta}_{\boldsymbol{k}})$ is a matrix which is a function of the hidden state $\boldsymbol{\theta}_k$, $\mathbf{a}_k$ is a vector of amplitudes such that the observation can be represented as a linear combination of basis functions (given by the columns of $\mathbf{G}_k$), $\mathbf{e}_k \sim \mathcal{N}(\mathbf{0}, \sigma^2_{(e),k}\mathbb{I}_L)$ is the observation noise vector and the length of the vector $\mathbf{x}_k$ is given by L.

In the context of the frequency estimation problem addressed in this thesis, $\boldsymbol{\theta}_k$ represents the vector of unknown frequencies present in the signal $\mathbf{x}_k$. The matrix $\mathbf{G}_k$ consists of cosine and sine functions of the frequencies over the duration of the signal, and the vector $\mathbf{a}_k$ denotes the associated amplitudes of the cosine and sine components of each frequency. This is discussed in more depth in Section 4.4 and the reader is referred to that section for further details of the model.

The state space equations do not describe the behaviour of the amplitude vector $\mathbf{a}_k$. Similar to the MPF derivation in Section 5.1, the amplitudes are assumed independent over time. Addi-

tionally, the vector $\mathbf{a}_k$ is treated as *multiplicative noise* and is assumed distributed according to a multivariate normal distribution $\mathbf{a}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_k)$, $\mathbf{\Sigma}_k = \sigma^2_{(a),k}\mathbb{I}_L$.

Implementation of the particle filter requires an expression for the likelihood function which will be derived next. Both the amplitude and noise vectors are assumed to follow normal distributions and consequently, the observation vector will also possess a normal distribution for a given $\boldsymbol{\theta}$ (and consequently $\mathbf{G}_k$). Multiplying $\mathbf{G}_k$ and $\mathbf{a}_k$ yields the result:

$$\mathbf{G}_k\mathbf{a}_k \sim \mathcal{N}(\mathbf{0}, \sigma^2_{(e),k}\delta^2_k\mathbf{G}_k\mathbf{G}^T_k). \tag{A.3}$$

The observation vector is then distributed according to the sum of two normal distributions and is given by:

$$\mathbf{x}_k \sim \mathcal{N}(\mathbf{0}, \sigma^2_{(a),k}\mathbf{G}_k\mathbf{G}^T_k + \sigma^2_{(e),k}) \tag{A.4}$$

and the likelihood function can be expressed as:

$$p(\mathbf{x}_k|\boldsymbol{\theta}_k, \sigma^2_{(a),k}, \sigma^2_{(e),k}) = \frac{1}{(2\pi)^{\frac{L}{2}}\sqrt{|\sigma^2_{(a),k}\mathbf{G}_k\mathbf{G}^T_k + \sigma^2_{(e),k}|}} \exp\left\{-\frac{1}{2\sigma^2_{(e),k}}\mathbf{x}^T_k(\sigma^2_{(a),k}\mathbf{G}_k\mathbf{G}^T_k + \sigma^2_{(e),k})^{-1}\mathbf{x}_k\right\}. \tag{A.5}$$

It is possible to implement a particle filter by using the above likelihood function to evaluate the importance weights. It is possible to marginalise the observation noise variance term from this distribution as was considered in Section 4.4.4 and this will be demonstrated next.

For convenience, the variance $\sigma^2_{(a),k}$ is parameterised in terms of the observation noise variance $\sigma^2_{(e),k}$ such that $\sigma^2_{(a),k} = \sigma^2_{(e),k}\delta^2_k$. This parameterisation is not strictly necessary, however, this aids in comparing the resulting filter in this section with that derived in Section 5.1.

An inverse gamma prior distribution is assumed for the observation noise variance $\sigma^2_{(e),k} \sim \mathfrak{IG}(\alpha_e, \beta_e)$ with scale and shape parameters (α_e, β_e) respectively. Using this prior distribution, it becomes possible to arrive at an expression for the conditional distribution $p(\mathbf{x}_k|\boldsymbol{\theta}_k, \delta^2_k)$ through the process of marginalisation:

$$p(\mathbf{x}_k|\boldsymbol{\theta}_k, \delta^2_k) = \int p(\mathbf{x}_k, \sigma^2_{(e),k}|\boldsymbol{\theta}_k, \delta^2_k)d\sigma^2_{(e),k} = \int p(\mathbf{x}_k|\boldsymbol{\theta}_k, \delta^2_k, \sigma^2_{(e),k})p(\sigma^2_{(e),k}|\alpha_e, \beta_e)d\sigma^2_{(e),k}. \tag{A.6}$$

The marginalised distribution is then given by:

$$p(\mathbf{x}_k|\boldsymbol{\theta}_k, \delta_k^2) = \frac{\alpha_e^{\beta_e}}{\Gamma(\beta_e)(2\pi)^{\frac{L}{2}}\sqrt{|\mathbf{F}_k|}} \int \sigma_{(e),k}^{-(\frac{L}{2}+\beta_e+1)} \exp\left\{-\frac{1}{2\sigma_{(e),k}^2}(\mathbf{x}_k^T\mathbf{F}_k^{-1}\mathbf{x}_k + 2\alpha_e)\right\} d\sigma_{(e),k}^2. \tag{A.7}$$

where $\mathbf{F}_k = \delta_k^2\mathbf{G}_k\mathbf{G}_k^T + \mathbb{I}_L$.

This integral can be performed by using the Gamma integral (see equation (4.39) on page 89). Setting $\alpha = \frac{1}{2}(\mathbf{x}_k^T\mathbf{F}_k\mathbf{x}_k + 2\alpha_e)$ and $\beta = \frac{L}{2} + \beta_e$ in the Gamma integral allows marginalisation of the observation noise variance $\sigma_{(e),k}^2$ yielding the likelihood function:

$$p(\mathbf{x}|\boldsymbol{\theta}_k, \delta_k^2) = \frac{2^{\beta_e}\alpha_e^{\beta_e}\Gamma(\frac{L}{2}+\beta_e)}{\pi^{\frac{L}{2}}\Gamma(\beta_e)} \frac{(\mathbf{x}_k^T\mathbf{F}_k^{-1}\mathbf{x}_k + 2\alpha_e)^{-(\frac{L}{2}+\beta_e)}}{\sqrt{|\mathbf{F}_k|}}. \tag{A.8}$$

The above equation is similar to marginalised likelihood function given by equation (4.40) on page 89. The most notable difference is the lack of the term $\delta_k^{2O_k}$ in the denominator of the above expression, where the number of basis functions is given by $2O_k$. This term penalises the likelihood to a greater extent when a larger number of basis functions is used.

The likelihood functions derived here in equations (A.5) and (A.8) can be used to implement a particle filter as an alternative to that derived in Section 4.4.4. However, the additional penalty incurred due to $\delta_k^{2O_k}$ may prove beneficial in limiting the number of basis functions used.

Appendix B
Publications

The following papers, based on the work in this thesis have been published, and are reproduced in this appendix.

S. Nagappa and J. R. Hopgood, *Frequency Tracking of Biological Waveforms*, in Proceedings of the IMA International Conference on Mathematics in Signal Processing, Cirencester, UK, 2006.

S. Nagappa and J. R. Hopgood, *Frequency Tracking using Monte Carlo Methods: Application to Bat Echolocation Signals*, in 16th Proceedings of the European Signal Processing Conference, Lausanne, Switzerland, August, 2008.

S. Nagappa and J. R. Hopgood, *A Sequential Monte Carlo Analysis Method for Bat Biosonar Signals*, in BioAcoustics 2009, Loughborough, UK, April 2009.

Frequency tracking of biological waveforms

By Sharad Nagappa, James R. Hopgood †

Institute for Digital Communications, Alexander Graham Bell Building
University of Edinburgh, Edinburgh - EH9 3JL
S.Nagappa@ed.ac.uk, James.Hopgood@ed.ac.uk

Abstract

This paper describes a particle filtering method for tracking the frequency of a mono-component signal. The method is aimed at tracking echolocation signals used by micro-bats which commonly have a single dominant harmonic component. In the approach followed here, the frequency of the mono-component signal is modelled as a slowly varying dynamic state. The particle filter is then used to track the frequency of this signal.

1. Introduction

Bats use a variety of echolocation calls. These calls are not only species-dependent, but also situation-specific. While many calls may be described as linear or exponential chirps, some are more complicated and cannot be represented in such simplistic terms. A further complication of the call structure is the presence of higher-order partial overtones. These characteristics can make a model based approach difficult.

One way of examining bat signals is to extract their time-varying parameters. Dubois *et al* (2005) use a particle filtering algorithm to track multiple spectral components of a signal using the short-time Fourier transform (STFT). In this paper, we attempt to track the frequency of a signal using a time-domain formulation. The method operates on the raw data, without any pre-processing, and therefore utilises all the information available in the data. The signal is constrained to a mono-component signal (or one with a single, highly-dominant frequency component). In addition, we are not interested in the estimates of amplitude or phase, and these are marginalised out of the likelihood. We provide some comparison of results with the method developed by Dubois *et al* which tracks frequencies using the STFT.

A simple harmonic frequency is modelled using the equation (Bretthorst (1988))

$$g(t) = B_1 \cos(2\pi f t) + B_2 \sin(2\pi f t) \tag{1.1}$$

where the parameters B_1, B_2 contain the amplitude and phase, and f is the frequency. Expressed in polar coordinates, this is identical to the formulation

$$g(t) = B\cos(2\pi f t + \phi); \quad \text{where} \quad B = \sqrt{B_1^2 + B_2^2}\ ; \quad \phi = \tan^{-1}\left\{\frac{B_2}{B_1}\right\}$$

If $n(t)$ is white Gaussian observation noise, the observation equation can be written as

$$z(t) = g(t) + n(t), \ \text{where } n(t) \sim \mathcal{N}(0, \sigma_n^2) \tag{1.2}$$

We define $\mathbf{z}_t = \{z(t-L+1), \cdots, z(t)\}$ as a vector of observations of length L. From

† This work was sponsored by the BIAS consortium under a grant provided by the EPSRC under the Basic Technology Programme. Bat data courtesy of Dr. Waters, U. of Leeds.

equation 1.2, the likelihood function is obtained as

$$p(\mathbf{z}_t \mid B_1, B_2, f, \sigma_n^2) \propto \prod_{i=t-L+1}^{t} \left[\frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left\{ -\frac{[z(i) - g(i)]^2}{2\sigma_n^2} \right\} \right]$$
$$\propto \sigma_n^{-L} \exp\left\{ -\frac{1}{2\sigma_n^2} \sum_{i=t-L+1}^{t} [z(i) - g(i)]^2 \right\} \tag{1.3}$$

The joint posterior distribution obtained from equation 1.3 is

$$p(B_1, B_2, f, \sigma_n^2 \mid \mathbf{z}_t) \propto p(\mathbf{z}_t \mid B_1, B_2, f, \sigma_n^2) \tag{1.4}$$

The distribution of interest is $p(f \mid z_k, \sigma_n^2)$. This may be obtained by marginalising the nuisance parameters B_1 and B_2 which constitute the amplitude and phase of the signal.

2. Marginalisation of nuisance parameters

Bretthorst (1988) illustrates the marginalisation of the amplitude-phase components for a single sinusoid modelled using equation 1.1. An uninformative uniform prior is assumed for these components. Bretthorst makes the assumption that the length of the observation vector is large, $L \gg 1$, in order to derive the marginalised distribution. Integrating out B_1 and B_2 then yields the joint likelihood

$$p(\mathbf{z}_t \mid f, \sigma_n^2) \propto \sigma_n^{-L+2} \exp\left\{ -\frac{L}{2\sigma_n^2} \left[\overline{\mathbf{z}_t^2} - \frac{2C(f)}{L} \right] \right\} \tag{2.1}$$

where $$\overline{\mathbf{z}_t^2} = \frac{1}{L} \sum_{i=t-L+1}^{t} z(i)^2$$

and $C(f)$, is the Schuster periodogram defined as

$$C(f) = \frac{1}{L} \left[\left\{ \sum_{i=t-L+1}^{t} z(i) \cos(2\pi f i) \right\}^2 + \left\{ \sum_{i=t-L+1}^{t} z(i) \sin(2\pi f i) \right\}^2 \right]$$

Using this formulation, the periodogram appears in a more natural way in the likelihood function, as compared to the STFT-method.

3. Particle filter framework

The particle filter relies on the transition probability and likelihood function which may be derived from the state-space equations shown below.

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{v}_{k-1} \tag{3.1}$$
$$\mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{n}_k \tag{3.2}$$

where $\mathbf{x}_k$ and $\mathbf{z}_k$ represent the state and observation vectors respectively at time k, $f(\cdot)$ and $h(\cdot)$ are (non-) linear functions. The process noise, $\mathbf{v}_{k-1}$, and the observation noise, $\mathbf{n}_k$, are assumed known.

The transition probability, $p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$, is used to update the process state, and the the posterior distribution of interest is proportional to the likelihood

$$p(\mathbf{x}_k \mid \mathbf{z}_k) \propto p(\mathbf{z}_k \mid \mathbf{x}_k) \tag{3.3}$$

4. Particle filtering for frequency tracking

Let f_k denote a time-varying frequency at time k. The frequency is modelled as a random walk with variance σ_f^2, under the constraint $0 < f_k < 0.5$ assuming a normalised sampling frequency of 1Hz. The process equation is written as

$$f_k = f_{k-1} + v_{k-1} \tag{4.1}$$

where $v_{k-1} \sim \mathcal{N}(0, \sigma_f^2)$. Equation 4.1, provides the transition probability $p(f_k \mid f_{k-1})$.

The observation z_k, at time k, is obtained using equation 1.2

$$z_k = B_{1,k}\cos(2\pi f_k k) + B_{2,k}\sin(2\pi f_k k) + n_k \tag{4.2}$$

where $B_{1,k}$ and $B_{2,k}$ denote (time-varying) amplitudes at time k; the observation noise is assumed to be Gaussian, $n_k \sim \mathcal{N}(0, \sigma_n^2)$, with σ_n^2 known.

We define $\mathbf{z}_k = \{z_{k-L+1}, \cdots, z_k\}$ as the observation vector; the signal is assumed to be slowly varying so that $f_{k-L+1} \approx f_k$, $B_{1,k-L+1} \approx B_{1,k}$, $B_{2,k-L+1} \approx B_{2,k}$. Under these assumptions, and for $L \gg 1$, $p(f_k \mid \mathbf{z}_k) \propto p(\mathbf{z}_k \mid f_k)$, defined in equation 2.1.

Algorithm 1 illustrates how the particle filter is used for frequency tracking. f_k^i denotes the frequency indicated by the i^{th} particle at time k. N is the number of particles used; $\widehat{N}_{eff}$ is the effective number of particles as defined by Ristic *et al* (2004), and N_{thr} is the threshold below which the resampling operation is performed.

Algorithm 1 SIR Particle Filter for Frequency Tracking

(*a*) FOR $i = 1 : N$
- Draw $f_k^i \sim p\left(f_k \mid f_{k-1}^i\right)$
- Evaluate the importance weights $\tilde{w}_k^i = w_{k-1}^i p\left(\mathbf{z}_k \mid f_k^i\right)$

(*b*) Calculate total weight $t = \text{SUM}\left[\{\tilde{w}_k^i\}_{i=1}^N\right]$

(*c*) FOR $i = 1 : N$
- Normalise $w_k^i = t^{-1}\tilde{w}_k^i$

(*d*) Calculate $\widehat{N}_{eff}$; IF $\widehat{N}_{eff} < N_{thr}$
- Resample: $\left[\{f_k^i, w_k^i\}_{i=1}^N\right] = \text{RESAMPLE}\left[\{f_{k-1}^i, w_{k-1}^i\}_{i=1}^N\right]$

5. Results

We provide results from simulations while varying three parameters: SNR of the signal, window-length, and number of particles used. Errors are provided as a relative percentage error. The test signal used is a linear chirp, and the error is shown as a function of the actual signal frequency rather than time. This can provides a measure of the frequency-dependent estimation error on the window-length L.

Figure 1 shows the frequency content of a synthetic test signal, with the corresponding STFT, at an SNR of 12dB, overlayed with the estimated frequency from a test run.

Figure 3 shows the error when the window-length, L, is varied at a constant SNR. The error decreases with increase in window-length, however, for a given window length, the error is higher when tracking a low-frequency signal.

In figure 4, the error obtained for different SNR conditions is compared, providing a limiting condition for satisfactory performance of the filter. At low SNR ($<$ 0dB), the error increases significantly; the error can be reduced by increasing the window length, however, this may violate the assumption of a slowly varying signal. The number of particles used in the simulations was fixed at 500 since there was no noticeable reduction in the error when more particles were used.

For comparison purposes, figure 4 also shows the estimation error obtained from ap-

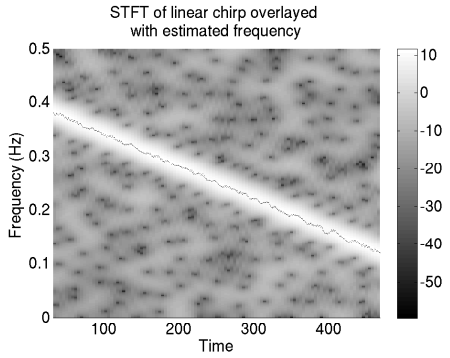


FIGURE 1. Tracking a linear chirp

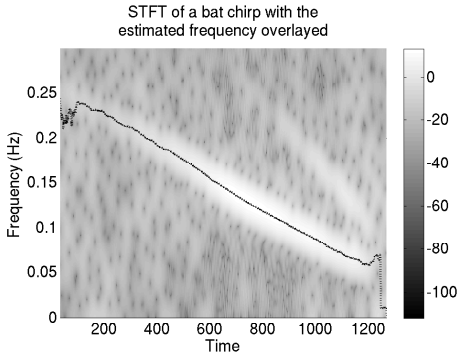


FIGURE 2. Tracking a linear bat chirp

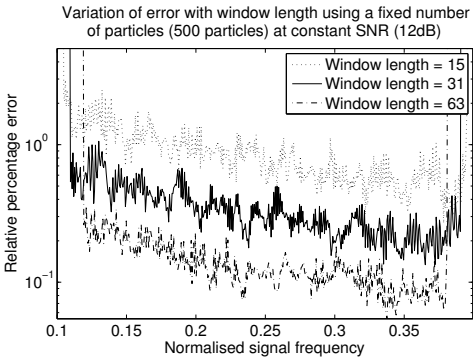


FIGURE 3. Error vs. window length

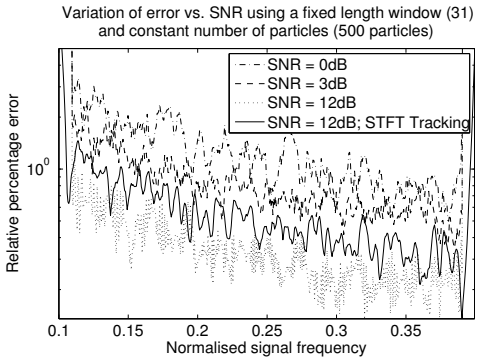


FIGURE 4. Error vs. SNR

plying a filter based on the method of Dubois *et al* (2005) to track the spectrum (labelled 'STFT tracking'). From the figure, our method seems to perform better. Further, compared to the method of Dubois *et al*, the formulation here is straightforward and uncomplicated, identical to the SIR particle filter, with the exception of the marginalisation, which is needed to obtain the likelihood.

A sample recording of a bat call from the bat *Myotis Nattereri* was used as a final test sequence to measure the robustness of the algorithm to real-world data. The result, shown in figure 2, appears to be satisfactory.

6. Conclusions

The method developed here offers a simple method for frequency-tracking of monocomponent signals using a particle filter. Results indicate that the method is both, reliable as well as robust. Work is currently under way to extend the method to deal with multi-component signals with a time-varying number of components.

REFERENCES

BRETTHORST, G. L. 1988 Bayesian Spectrum Analysis and Parameter Estimation. *Springer-Verlag; Lecture Notes in Statistics*.

DUBOIS, C.; DAVY, M. & IDIER, J. 2005 Tracking of Time-Frequency Components using Particle Filtering. *Proceedings. (ICASSP '05)*.

RISTIC, B.; ARULAMPALAM, S. & GORDON, N. 2004 Beyond the Kalman Filter: Particle Filters for Tracking Applications. *Artech House Publishers*.

FREQUENCY TRACKING USING MONTE CARLO METHODS: APPLICATION TO BAT ECHOLOCATION SIGNALS

Sharad Nagappa and James Hopgood

Institute for Digital Communications, University of Edinburgh,
Edinburgh - EH9 3JL, United Kingdom
email: S.Nagappa@ed.ac.uk, James.Hopgood@ed.ac.uk

ABSTRACT

This paper presents a sequential Monte Carlo method for tracking an unknown varying number of time-varying frequencies. A reversible-jump sampler is used to implement model-order determination. It is shown that for a linear-in-the-amplitudes observation model in white Gaussian noise, the amplitudes and noise variance can be analytically marginalised out of the posterior distribution resulting in a reduced dimension state estimation problem. A sum of linear chirps model is chosen as a local observation model and this basis is used to determine the instantaneous frequency of the signal. We present frequency tracking results from synthetic as well as field-recorded bat echolocation calls and compare results with Fourier based frequency tracking.

1. INTRODUCTION

Time frequency analysis of biosonar signals can be performed using techniques based around the short-time Fourier transform (STFT) or wavelets, with further analysis relying on this time-frequency representation. This work addresses the issue of instantaneous frequency estimation by direct inference from the raw data and without any preprocessing. A sequential Bayesian approach is adopted to solve this problem. The use of a Bayesian approach facilitates incorporation of prior information into the estimator, allowing the estimator to use all the information available.

Figure 1 [1] illustrates the time-varying signal structure as a bat attempts to first locate and identify potential targets (usually insects, as a source of food), then tracks the target until it is finally captured. In the figure, there are certain obvious changes in the signal structure over the *search-approach-terminal* phases: the duration of each call decreases, multiple harmonics may be introduced, and the pulse repetition rate increases. There is very little quantitative analysis of the bat calls in the literature that exists beyond such a descriptive analysis. The aim of this work is to then provide a robust method to determine and track the time-varying frequency of multiple components of such a signal, thus providing a basis for quantitative analysis of these signals. Such an analysis would be useful to biologists studying echolocation in bats.

The implementation of a sequential approach offers several advantages over a batch approach. A batch approach necessitates a model for the time-frequency structure of the bat call which varies across species. In addition, each call can have different amplitude envelopes, and multiple harmonics, which may not last the entire duration of the signal. The batch approach can result in a difficult estimation problem with many parameters.

To overcome the problem of "parameter-bloat", we choose to sequentially estimate the frequencies of the subcomponents present in the signal. As a result, no specific model needs to be considered for different species while estimating the frequency. Using these frequency estimates, model determination and parameter estimation can be performed subsequently.

IDCOM is in the Joint Research Institute for Signal and Image Processing, a member of the Edinburgh Research Partnership in Engineering and Mathematics. This work is sponsored by the EPSRC under their basic technology grant, "Biologically Inspired Acoustic Systems". Bat recordings used in this work were obtained from Dr Dean Waters, University of Leeds.

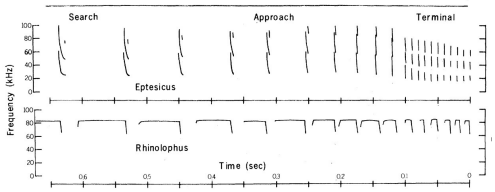


Figure 1: Diversity of echolocation call structure

One approach for estimating the instantaneous frequency content of a signal would be to locate peaks in the spectrogram of a signal. Such a simplistic approach is not necessarily either reliable or robust. Approaches for time-varying frequency tracking have been proposed previously [2, 3, 4] taking into account smooth changes in the frequency trajectory. STFT-based time-frequency tracking [2] is implemented by applying a particle filter [5] to track the peaks present in the spectrogram of a signal. The method looks at the Fourier transform over short blocks of the observation sequence and estimates in each block the number of components present, as well as their frequencies, amplitudes, and noise variance.

The method described here is related to the harmonic tracking algorithm described in [4]. In their work, a jump Markov system (JMS) is used to detect multiple harmonic components. A Rao-Blackwellised particle filter (RBPF) [5] is used to integrate out amplitude parameters in their model. In contrast, we investigate an alternative approach where the amplitude and noise variance parameters are analytically marginalised out due to the use of a linear-in-the-amplitudes observation model in Gaussian noise [6, 7]. Further, reflecting the data set of interest, viz., bat echolocation calls, we adopt a linear chirp as opposed to a sinusoidal basis.

When dealing with biosonar signals, the number of components present in the signal is a time-varying parameter. Reversible-jump Markov chain Monte Carlo (RJMCMC) methods [8] have previously been used to estimate an unknown number of frequencies for the stationary-frequency case [9]. Within a sequential Monte Carlo (SMC) framework, we demonstrate that the idea can be extended and applied to the problem of frequency-tracking when the number of frequency components is time-varying. While a JMS can be applied to detect a varying number of components, the RJMCMC sampler is incorporated into a SMC framework since it improves filter performance by inhibiting unlikely moves [10].

2. SIGNAL MODEL

The rate of change of phase of a signal provides the instantaneous frequency of the signal. For a multicomponent signal, it is not the overall rate of change of change of phase that we are interested in, but the combination of frequencies which are present at that instant. We redefine "instantaneous frequency" here so that each subcomponent of the signal has its own instantaneous frequency.

In order to estimate the instantaneous frequencies, we slide a window over the observation sequence and estimate the frequency

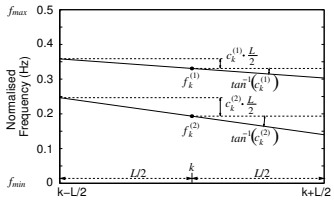


Figure 2: Over the sliding window, the signal is modelled as the sum of linear chirps. In this case, there are two linear chirps with centre frequencies $(f_k^{(1)}, f_k^{(2)})$, and chirp-rates $(c_k^{(1)}, c_k^{(2)})$.

components present in the windowed section of the observations. The time varying nature of each frequency component is modelled along the lines of the equation

$$f_{k+1} = g(f_k) + w_k \tag{1}$$

where f_k is the frequency at time k, $g(\cdot)$ is a function to update the frequency from the previous time instant, and $w_k \sim \mathcal{N}(0, \sigma_w^2)$.

Within the sliding window, the observation segment is modelled locally as the sum of P_k linear chirps with frequencies $\mathbf{f}_k = [f_k^{(1)}, f_k^{(2)}, \ldots, f_k^{(P_k)}]^T$ and chirp-rates $\mathbf{c}_k = [c_k^{(1)}, c_k^{(2)}, \ldots, c_k^{(P_k)}]^T$. Let L be the length of the sliding window defined such that k is at the centre of the window. Over this window, the observation x_t at time $t \in \{k-L/2, \ldots, k+L/2\}$, assuming L odd, is

$$x_t = \sum_{p=1}^{P_k} \left(a_k^{(p)} \cos \phi_{k,t}^{(p)} + b_k^{(p)} \sin \phi_{k,t}^{(p)} \right) + n_t \tag{2}$$

$$\phi_{k,t}^{(p)} = 2\pi \left(f_k^{(p)} - c_k^{(p)} \frac{L}{2} + \frac{c_k^{(p)}}{2} t \right) t \tag{3}$$

where $\mathbf{a}_k = [a_k^{(1)}, a_k^{(2)}, \ldots, a_k^{(P_k)}, b_k^{(1)}, b_k^{(2)}, \ldots, b_k^{(P_k)}]^T$ are the amplitudes of the cosine and sine amplitudes of the subcomponents, and $n_t \sim \mathcal{N}(0, \sigma_{n,k}^2)$ is zero-mean white Gaussian observation noise with variance $\sigma_{n,k}^2$. Setting $c_k^{(p)}$ to zero in equation 3 reduces the local model to a sum of sinusoids (SoS) model. Figure 2 illustrates the parameters used in the local model.

The selection of a suitable window length depends on the application under consideration. Shorter windows allow us to deal with highly non-stationary frequencies since the variation across a short window will be less. However, the short window also leads to greater variance in the parameter estimate. Long windows, by contrast, yield lower variance estimates, but may yield worse estimates due to frequency non-stationarity within the window. This trade-off must be considered when selecting a window of suitable length.

3. THE LIKELIHOOD FUNCTION

Since the noise term in equation 2 is zero-mean white Gaussian noise, the likelihood function can easily be written down. The likelihood function is evaluated over the windowed section of the observations. Let $\mathbf{x}_k = [x_{k-L/2}, \ldots, x_{k+L/2}]^T$ be the vector of observations; $\mathbf{n}_k = [n_{k-L/2}, \ldots, n_{k+L/2}]^T$ be the noise.

The model in equation 2 can be written in matrix vector form by defining the following matrices:

$$\mathbf{\Phi}_k = \begin{bmatrix} \phi_{k,-\frac{L}{2}}^{(1)} & \phi_{k,-\frac{L}{2}}^{(2)} & \cdots & \phi_{k,-\frac{L}{2}}^{(P_k)} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{k,\frac{L}{2}}^{(1)} & \phi_{k,\frac{L}{2}}^{(2)} & \cdots & \phi_{k,\frac{L}{2}}^{(P_k)} \end{bmatrix}$$

$$\mathbf{G}_k = [\cos(\mathbf{\Phi}_k) \;\; \sin(\mathbf{\Phi}_k)]$$

where $\phi_{k,t}$ is defined in equation 3; $\cos(\cdot)$ and $\sin(\cdot)$ operate element-wise on the matrix $\mathbf{\Phi}_k$ to produce an augmented $L \times 2P_k$ $\mathbf{G}_k$ matrix. The signal model can be rewritten in the form of the general linear model (GLM) [6, 7] as

$$\mathbf{x}_k = \mathbf{G}_k \mathbf{a}_k + \mathbf{n}_k \tag{4}$$

If $\psi_k = \{P_k, \mathbf{f}_k, \mathbf{c}_k, \mathbf{a}_k, \sigma_{n,k}^2\}$, the likelihood function is

$$p(\mathbf{x}_k \mid \psi_k) = \frac{1}{(2\pi\sigma_{n,k}^2)^{\frac{L}{2}}} \exp\left[-\frac{\|\mathbf{x}_k - \mathbf{G}_k \mathbf{a}_k\|^2}{2\sigma_{n,k}^2} \right] \tag{5}$$

3.1 Parameter Reduction using Marginalisation

From equation 5, it is possible to obtain a likelihood distribution with the amplitude and noise variance terms marginalised out [7]. We would like to remove the dependence of equation 5 on $\Upsilon_k = \{\mathbf{a}_k, \sigma_{n,k}^2\}$, thus reducing the parameter space. The frequency and chirp-rate parameters can then be estimated, for example, using a non-linear search. Since the observations are linear in the amplitude parameters, these can be estimated separately once the frequencies and chirp-rates are determined.

A multivariate normal distribution is defined on the $2P_k$ amplitude parameters, $\mathbf{a}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_k), \mathbf{\Sigma}_k = \sigma_{n,k}^2 \delta_k^2 \mathbb{I}_{2P_k}$, where $\mathbb{I}_Q$ is the $Q \times Q$ identity matrix. An extra parameter, δ_k^2, is introduced, which is indicative of the SNR of the signal. The noise variance is assumed to follow an inverse-gamma distribution, $\sigma_{n,k}^2 \sim \mathcal{IG}(\alpha_n, \beta_n)$, with scale and shape parameters (α_n, β_n). Thus, more explicitly,

$$p(\mathbf{a}_k \mid P_k, \sigma_{n,k}^2, \delta_k^2) = \frac{1}{(2\pi\delta_k^2\sigma_{n,k}^2)^{P_k}} \exp\left(-\frac{1}{2} \mathbf{a}_k^T \Sigma_k^{-1} \mathbf{a}_k \right)$$

$$p(\sigma_{n,k}^2 \mid \alpha_n, \beta_n) = \frac{\alpha_n^{\beta_n}}{\Gamma(\beta_n)} (\sigma_{n,k}^2)^{-(\beta_n+1)} \exp\left(-\frac{\alpha_n}{\sigma_{n,k}^2} \right)$$

The marginalised likelihood function, $p(\mathbf{x}_k \mid \psi_k')$, where $\psi_k' = \{P_k, \mathbf{f}_k, \mathbf{c}_k, \delta_k^2\}$, is then obtained from:

$$p(\mathbf{x}_k \mid \psi_k') = \int p(\mathbf{x}_k, \Upsilon_k \mid \psi_k') d\Upsilon_k = \int p(\mathbf{x}_k \mid \psi_k', \Upsilon_k) p(\Upsilon_k) d\Upsilon_k$$

$$\propto \frac{\left(\mathbf{x}_k^T \mathbf{x}_k + 2\alpha_n - \mathbf{x}_k^T \mathbf{G}_k \mathbf{F}_k^{-1} \mathbf{G}_k^T \mathbf{x}_k \right)^{-(\frac{L}{2}+\beta_n)}}{\delta_k^{2P_k} \sqrt{\det(\mathbf{F}_k)}} \tag{6}$$

where $\mathbf{F}_k = \mathbf{G}_k^T \mathbf{G}_k + \delta_k^{-2} \mathbb{I}_{2P_k}$.

Priors on the remaining parameters are specified as uniform on the frequency and chirp rate, an inverse-gamma on δ_k^2, $p(\delta_k^2) = \mathcal{IG}(\alpha_\delta, \beta_\delta)$, and a Poisson distribution on P_k, $p(P_k) = p(P_k \mid \lambda)$.

3.2 Parameter Estimation

To obtain, for example, the maximum likelihood estimate, a difficult multidimensional search for the function maximum is necessary. This is further complicated by the fact that the number of parameters to be estimated is a parameter that needs to estimated as well. The complexity can be reduced by assuming that the signal does not change significantly between time k and $k+1$ and exploiting this within a sequential Bayesian framework. We adopt this solution to estimate the signal parameters by modelling the sequential update of the parameters along the lines of equation 1.

Particle filtering methods offer a framework for the implementation of the recursive Bayesian filter, and we use this framework to estimate and track the frequency content of the time-varying signal.

4. PARTICLE FILTERING FRAMEWORK

A sequential importance-sampling resampling (SIR) particle filter [5] is used to perform online frequency-tracking. The particle filter would normally be used to estimate all the state parameters. However, we will demonstrate next that for the GLM, within a sequential framework, the amplitude and noise variance terms may be analytically marginalised out of the posterior distribution.

Since many problems can be placed within the context of the GLM, this marginalisation has wide application. The benefit is that

while the marginalisation introduces an extra parameter, at the same time $2P_k+1$ parameters are eliminated from the likelihood function.

A particle filter approximates the posterior distribution by a set of weighted samples using sequential importance-sampling (SIS). Let ω_k be the unknown state at time k, $\Omega_k = \{\omega_j\}_{j=0}^{k}$, and $\mathbf{x}_k$ be the observation, $\mathbf{X}_k = \{\mathbf{x}_j\}_{j=0}^{k}$. Then $p(\Omega_k \mid \mathbf{X}_k)$, the posterior distribution, is approximated by a set of N discrete weighted samples, or particles, $\{\Omega_k^{(i)}, w_k^{(i)}\}_{i=1}^{N}$, where $w_k^{(i)}$ is the weight associated with the sample $\mathbf{\Omega}_k^{(i)}$.

$$p(\Omega_k \mid \mathbf{X}_k) \approx \sum_{i=1}^{N} w_k^{(i)} \delta(\Omega_k - \Omega_k^{(i)}) \tag{7}$$

The weight $w_k^{(i)}$ evaluates to

$$w_k^{(i)} \propto \frac{p(\mathbf{x}_k \mid \omega_k^{(i)}) p(\omega_k^{(i)} \mid \omega_{k-1}^{(i)}) p(\Omega_{k-1}^{(i)} \mid \mathbf{X}_{k-1})}{q(\omega_k^{(i)} \mid \Omega_{k-1}^{(i)}, \mathbf{X}_k) q(\Omega_{k-1}^{(i)} \mid \mathbf{X}_{k-1})} \tag{8}$$

$$= w_{k-1}^{(i)} \frac{p(\mathbf{x}_k \mid \omega_k^{(i)}) p(\omega_k^{(i)} \mid \omega_{k-1}^{(i)})}{q(\omega_k^{(i)} \mid \Omega_{k-1}^{(i)}, \mathbf{X}_k)} \tag{9}$$

where $w_{k-1}^{(i)}$ is the weight at the previous instant $k-1$, the likelihood $p(\mathbf{x}_k \mid \omega_k^{(i)})$ is obtained from equation 5, $p(\omega_k^{(i)} \mid \omega_{k-1}^{(i)})$, the transition prior, is obtained according to the state update equations (section 4.1). $q(\omega_k^{(i)} \mid \Omega_{k-1}^{(i)}, \mathbf{X}_k)$ is the sampling distribution dependent on previous states and observations.

To marginalise the parameters $\Upsilon_k = \{\mathbf{a}_k, \sigma_{n,k}^2\}$ from the posterior, we must integrate over these parameters in equation 7. It is possible to use Rao-Blackwellisation to marginalise the linear amplitude parameters [4], however, we adopt a different approach in analytically marginalising the amplitude and noise variance parameters. Rather than assume a particular model on the Υ_k parameters (e.g. random walk, as in [4]), we assume that the parameters are independent across windows which allows us to easily carry out the marginalisation. The marginalised posterior is written as:

$$p(\Omega'_k \mid \mathbf{X}_k) \approx \int \cdots \int \sum_{i=1}^{N} w_k^{(i)} \delta(\Omega'_k - \Omega'^{(i)}_k) d\Upsilon_{0:k} \tag{10}$$

$$\approx \sum_{i=1}^{N} w'^{(i)}_k \delta(\Omega'_k - \Omega'^{(i)}_k)$$

where $\Omega'_k = \Omega_{k-\{\Upsilon_{0:k}\}}$ and $w'_k = \int \cdots \int w_k d\Upsilon_{0:k}$.

From equation 10 the posterior distribution can be obtained by the integral of the weight in equation 8 with respect to the nuisance parameters giving rise to a 'marginalised' weight $w'^{(i)}_k$. The derivation of this marginalised weight update is listed in appendix A.

We utilise the priors listed in section 3.1 and set $\omega'_k = \psi'_k = \{P_k, \mathbf{f}_k, \mathbf{c}_k, \delta_k^2\}$. Under the assumption that $p(\Upsilon_k \mid \Upsilon_{k-1}) = p(\Upsilon_k)$, the marginalised likelihood $p(\mathbf{x}_k \mid \omega'_k)$ is identical to equation 6. In addition, the sampling distribution in the SIS step is chosen as the transition prior resulting in the simplified weight update equation

$$w'^{(i)}_k \propto w'^{(i)}_{k-1} p(\mathbf{x}_k \mid \psi'^{(i)}_k) \tag{11}$$

In the approach adopted here, the SIR filter tracks frequencies which have already been detected but does not detect any change in the number of components. A RJMCMC move is subsequently introduced to detect changes in the number of signal components.

4.1 State Update Equations

The state parameters ψ' need to be updated from time t_k at index k to time $t_k + \Delta_t$ at $k+1$.

$$\begin{array}{rcccl} \mathbf{f}_{k+1} & = & \mathbf{f}_k + \mathbf{c}_k \Delta_t & + & \mathbf{v}_{f,k} \\ \mathbf{c}_{k+1} & = & \mathbf{c}_k & + & \mathbf{v}_{c,k} \\ \log(\delta_{k+1}^2) & = & \log(\delta_k^2) & + & v_{\delta,k} \end{array} \tag{12}$$

where $\mathbf{v}_{f,k} = [v_{f,k}^{(1)}, \ldots, v_{f,k}^{(P_k)}]^T$, $\mathbf{v}_{c,k} = [v_{c,k}^{(1)}, \ldots, v_{c,k}^{(P_k)}]^T$, $v_{f,k}^{(p)} \sim \mathcal{N}(0, \sigma_f^2)$, is the process noise for update of the p^{th} frequency, $v_{c,k}^{(p)} \sim \mathcal{N}(0, \sigma_c^2)$, is the process noise for update of the p^{th} chirp-rate, $v_{\delta,k} \sim \mathcal{N}(0, \sigma_\delta^2)$, is the process noise for update of the δ_k term. Since the signal is modelled as a linear chirp, $f_k^{(p)}$ can be expected to increase by the amount $\mathbf{c}_k \Delta_t$ at time index $k+1$, and the frequency update is formulated to mirror this.

Changes to the model order P_k are not reflected in the state space equations. The state space equations reflect the update behaviour of the SIR particle filter which is unable to cope with changes in the model order. The model order term is thus updated according to the RJMCMC sampler (section 4.3).

4.2 Instantaneous Parameter Estimates

The particle filter produces an approximation of the multidimensional distribution over the multiple signal parameters. Each particle contains an unordered set of frequency and chirp-rate parameters. As a result, taking a sample mean does not yield a meaningful estimate of the instantaneous frequency content of the signal. We choose the MAP estimate as being representative of the instantaneous signal parameters. The instantaneous estimate is written as

$$\hat{\psi}'_k = \arg_{\psi'_k} \max p(\psi'_k \mid \mathbf{X}_k) \tag{13}$$

Using the estimated ψ'_k parameters, it is possible to obtain an estimate of the amplitude parameters using a least squares solution:

$$\mathbf{a}_k = (\mathbf{G}_k^T \mathbf{G}_k)^{-1} \mathbf{G}_k^T \mathbf{x}_k$$

4.3 Model-Order Determination using Reversible-Jump Markov Chain Monte Carlo

An alternative to using JMS for model-order selection is to utilise a RJMCMC step in a particle filter which offers certain benefits. The reversible-jump sampler accepts a new state according to an acceptance probability. This accept/reject mechanism ensures that good particles are not lost, thus reducing the variance of the weights and limiting particle degeneracy [10]. This benefit, however, comes at the cost of applying the sampler to each particle in the filter.

In contrast with general MCMC methods, the reversible-jump sampler does not require a burn-in period when used in the SMC context. The reversible-jump sampler requires the burn-in period so as to sample from the limiting distribution. Since the samples from the SIR filter are already distributed accordingly, a single iteration is sufficient for model-order determination [11].

The reversible-jump move updates the number of components tracked by each particle at time k. Moves between different dimension spaces are performed using birth, death, and update moves [8, 9] with respective probabilities $\{b_k, d_k, u_k\}$. Let $\{P_k, \Theta_k\}$ denote the current state, and $\{P_k^\star, \Theta_k^\star\}$, the proposed state, where P_k is the number of components and $\Theta_k = \psi'_{k-\{P_k\}}$. The new state will be accepted according to an acceptance ratio r,

$$r = \underbrace{\frac{p(P_k^\star, \Theta_k^\star \mid \mathbf{x}_k, \Lambda)}{p(P_k, \Theta_k \mid \mathbf{x}_k, \Lambda)}}_{\text{posterior ratio}} \underbrace{\frac{d(P_k, \Theta_k \mid P_k^\star, \Theta_k^\star)}{d(P_k^\star, \Theta_k^\star \mid P_k, \Theta_k)}}_{\text{proposal ratio}} \underbrace{\mathbf{J}}_{\text{Jacobian}}$$

where $d(\cdot \mid \cdot)$ denotes the conditional proposal distribution for the parameters. The Jacobian term evaluates to unity for birth and death moves. Once a particular move type is selected, a new state is proposed which is then accepted with an acceptance probability $\alpha = \min\{1, r\}$. Algorithm 1 lists the implementation of the filter.

5. RESULTS

We first present a comparison of the SoS and SoLC models. This will illustrate the gains achieved from using the more complicated model. Since the target class of signals is bat echolocation calls

Algorithm 1 Tracking a multicomponent signal

Initialisation ($k=0$):
$(P_0^{(i)}, \mathbf{f}_0^{(i)}, \mathbf{c}_0^{(i)}, \delta_k^{2(i)}) \sim p(P, \mathbf{f}, \mathbf{c}, \delta^2) \mid_{i=1}^{N}$ where i is the particle index, and N is the number of particles used.

1. Increment $k \leftarrow k+1$
2. Process-update:
$$\begin{aligned} P_k^{(i)} &= P_{k-1} \mid_{i=1}^{N} \\ \mathbf{f}_k^{(i)} &\sim p(\mathbf{f} \mid P_{k-1}^{(i)}, \mathbf{f}_{k-1}^{(i)}, \mathbf{c}_{k-1}^{(i)}) \mid_{i=1}^{N} \\ \mathbf{c}_k^{(i)} &\sim p(\mathbf{c} \mid P_{k-1}^{(i)}, \mathbf{c}_{k-1}^{(i)}) \mid_{i=1}^{N} \\ \delta_k^{2(i)} &\sim p(\delta^2 \mid (\delta_{k-1}^2)^{(i)}) \mid_{i=1}^{N} \end{aligned}$$
3. Evaluate importance weight (equation 11):
$\hat{w}_k'^{(i)} \propto {w'_{k-1}}^{(i)} p(\mathbf{x}_k \mid \psi'_k)$
4. Normalise weights:
$w_k'^{(i)} = T^{-1} \hat{w}_k'^{(i)} \mid_{i=1}^{N}; \quad T = \sum_{i=1}^{N} \hat{w}_k'^{(i)}$
5. Obtain parameter estimates:
$\psi'_k = \arg_{\psi'_k} \max p(\psi' \mid \mathbf{X}_k, \Lambda)$
$\mathbf{a}_k = (\mathbf{G}_k^T \mathbf{G}_k)^{-1} \mathbf{G}_k^T \mathbf{x}_k$
6. Resample $[\psi_k'^{(i)}] = \text{RESAMPLE}[\{\psi_k'^{(i)}, w_k^{(i)}\}_{i=1}^{N}]$
7. Reversible jump move
For $i = 1 : N$
   - Sample $u \sim \mathcal{U}_{[0,1]}$; select move type according to probabilities b_k, d_k, u_k.
   - Propose new state $\{P_k^{\star(i)}, \Theta_k^{\star(i)}\}$ and evaluate acceptance probability α.
   - Sample $u \sim \mathcal{U}_{[0,1]}$; if $u \leq \alpha$, accept new state $\{P_k^{\star(i)}, \Theta_k^{\star(i)}\}$, otherwise retain old state $\{P_k^{(i)}, \Theta_k^{(i)}\}$.

which possess significant variety, the added flexibility of the SoLC model is useful. We will then show that the algorithm is able to detect changes in the number of subcomponents when using frequency modulated signals. We will finally present some results from tests on field-recorded bat echolocation calls.

5.1 SoS Model vs SoLC Model

As a test signal, we use a quadratic chirp having a constant-frequency (CF) tail (figure 3), at an SNR of approximately 12dB. Using a normalised sampling frequency of 1 Hz, a spectrogram of the signal (without noise) highlights the presence of multiple peaks (multiple dark tracks), which can lead to over-modelling of the signal. The number of components in the simulation is consequently constrained to 1, and 100 particles are used for all the simulations. The window length is set to 129 samples. 500 Monte Carlo runs were performed for each model and the average MSE was plotted.

A comparison of the average MSE shows that the SoLC model outperforms the SoS model for the FM section of the signal, and is only slightly worse in the CF tail. Also compared with these models, is the MSE from tracking peaks in the spectrogram. The method is similar, but not identical to the method of Dubois *et al* [2], relying implicitly on a SoS model, and is found to perform about as well as the SoS model used here. The use of a hamming window instead of a rectangular window, which is used in all the simulations here, can provide a slight improvement for the STFT tracker since this lowers the sidelobes in the Fourier transform.

The added complexity from the SoLC model allows us to obtain significantly better estimates from chirp signals. The presence of multiple peaks in the Fourier transform will result in over-modelling in the case of the SoS model, however, the SoLC model is flexible enough to deal with this non-stationarity.

5.2 Frequency Tracking Results

The purpose of these results is to illustrate filter performance at points where signals start and end. The algorithm is able to track multiple components, although there is some uncertainty in the estimate at the start and end of signals. At crossing points, the algorithm tracks the separate components and does not approximate them as a single component.

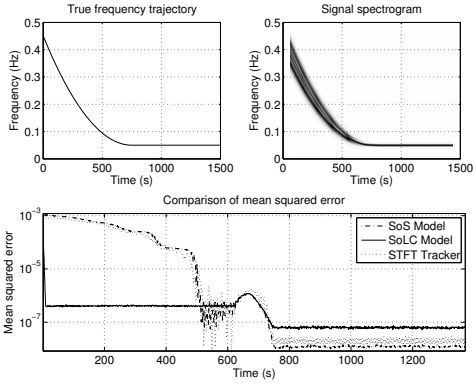


Figure 3: Comparison of SoS and SoLC models

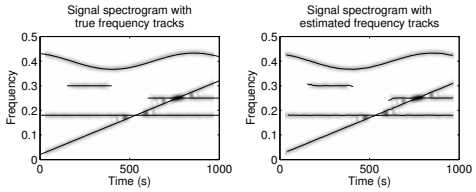


Figure 4: Test using multiple (and crossing) signals

The signal used is similar to that used in [3]. Figure 4 shows the spectrogram of the signal with the true frequency trajectories overoverlaid. The signal incorporates FM signals, multiple frequency crossings as well as a changing number of components, although none of the components is amplitude-modulated. The SNR of the signal is approximately 20 dB.

In the simulation, the number of particles used is 1000; a window-length of 65 samples is chosen, and the hyperparameters are set as $\alpha_n = \beta_n = 0$, $\alpha_\delta = 1000$, $\beta_\delta = 2$, $\lambda = 1$. The hyperparameters are deliberately chosen so as to specify vague priors on the parameters. Using $(\alpha_\delta = 1000, \beta_\delta = 2)$ specifies a prior distribution on the "SNR" term with mean 30dB.

The presence of discontinuities within the sliding window causes problems for the filter, which tends to over-model the signal in an attempt to fit the discontinuity. To overcome this problem, we assume that two frequency components will be no closer than a predefined limit. This is achieved in practice by inserting nulls into the sampling distribution for the frequency parameter. The disadvantage of this, however, is that frequencies which are very close together will not be detected using the reversible jump sampler.An alternate way of limiting this problem is to use a very short window.

5.3 Frequency Tracking Applied to Bat Calls

Results are shown here from testing the algorithm with two bat echolocation calls recorded in the field. 1000 particles were used and a window-length of 65 samples was specified. A truncated uniform prior is used for the distribution of the frequency component. Since these recordings contain a significant amount of low-frequency noise, the truncated prior is used to disregard the low-frequency noise band. The sampling distribution for adding new frequency components is chosen as a modified truncated prior containing nulls around already existing frequencies.

Figure 5 is a good quality recording of a Pipistrelle bat call (16 bits per sample (bps) with 750 kHz sampling rate). The signal has a single dominant component with another component approximately 30-dB below that. The algorithm is able to track the non-linear chirp as well as a higher harmonic. Figure 6 shows a recording from a

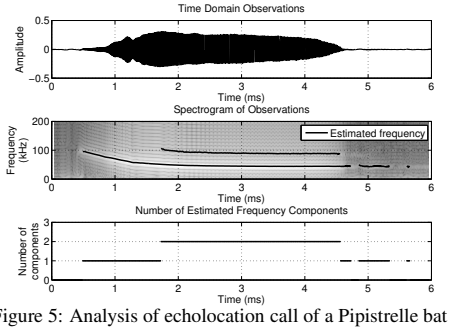


Figure 5: Analysis of echolocation call of a Pipistrelle bat

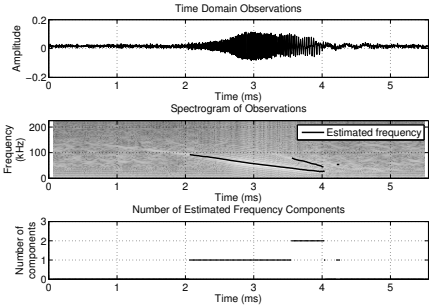


Figure 6: Analysis of echolocation call of a Natterer's bat

Natterer's bat with an SNR of 16 dB (8 bps with 450 kHz sampling rate). The two frequency components are detected in the signal.

The application of the algorithm to bat chirps allows us to extract a set of instantaneous frequencies from the signals. It is possible to transform these frequencies into individual signal tracks (for example, by applying target tracking algorithms). These tracks can then be studied to better describe the nature of bat calls in terms of linear/hyperbolic chirps with relevant parameters.

6. SUMMARY AND CONCLUSIONS

The method described here provides a means for detecting a time-varying number of dynamic frequencies by locally modelling the signal as a sum of linear chirps. The algorithm is based around direct inference from the observations without any form of preprocessing. The use of a sum of linear chirps basis is shown to offer advantages over a sum of sinusoids basis. A particle filtering framework allows frequency tracking in a sequential framework, while the RJMCMC sampler permits detection of the start/end of components. A marginalisation operation is used to reduce the dimension of the parameter state space. This marginalisation of amplitudes and noise variance is shown to be valid for the GLM framework.

A. MARGINALISATION FOR SEQUENTIAL UPDATE

The particle filter weight update equation is written as

$$w_k = w_{k-1} \frac{p(\mathbf{x}_k \mid \psi_k) p(\psi_k \mid \psi_{k-1})}{q(\psi_k \mid \Psi_{k-1}, \mathbf{X}_k)}$$

where w_{k-1} is the weight of a particle at the previous time instant $k-1$. If ψ'_k denotes the state parameters we wish to estimate, and Υ_k is the parameters we wish to marginalise out of the estimation, then the transition prior $p(\psi_k \mid \psi_{k-1})$ can be rewritten as

$$\begin{aligned} p(\psi_k \mid \psi_{k-1}) &= p(\psi'_k, \Upsilon_k \mid \psi'_{k-1}, \Upsilon_{k-1}) \\ &= p(\psi'_k \mid \psi'_{k-1}, \Upsilon_k, \Upsilon_{k-1}) p(\Upsilon_k \mid \Upsilon_{k-1}, \psi'_{k-1}) \\ &= p(\psi'_k \mid \psi'_{k-1}) p(\Upsilon_k \mid \Upsilon_{k-1}) \end{aligned}$$

We make a further approximation that the Υ parameters are completely independent across blocks, such that $p(\Upsilon_k \mid \Upsilon_{k-1}) = p(\Upsilon_k)$, where $p(\Upsilon_k)$ represents some prior distribution. This allows us to carry out the marginalisation with relative ease by removing any dependence on previous states.

The sampling distribution reflects the parameters being drawn and is specified as $q(\psi_k \mid \Psi_{k-1}, \mathbf{X}_k) = q(\psi'_k \mid \psi'_{k-1}, \mathbf{x}_k)$, i.e., dependent on parameters of interest and the most recent observation.

Evaluation of the weight w'_k is key to computation of the marginalised posterior distribution.

$$w'_k = \int \cdots \int \frac{p(\mathbf{x}_k \mid \psi_k) p(\psi_k \mid \psi_{k-1}) p(\Psi_{k-1} \mid \mathbf{X}_{k-1})}{q(\psi_k \mid \Psi_{k-1}, \mathbf{X}_k) q(\Psi_{k-1} \mid \mathbf{X}_{k-1})} d\Upsilon_{0:k}$$

The marginalised weight evaluates to:

$$\begin{aligned} w'_k &\propto \frac{p(\psi'_k \mid \psi'_{k-1})}{q(\psi'_k \mid \psi'_{k-1}, \mathbf{x}_k)} \int p(\mathbf{x}_k \mid \psi_k) p(\Upsilon_k) d\Upsilon_k \\ &\quad \times \int \cdots \int p(\Psi_{k-1} \mid \mathbf{X}_{k-1}) d\Upsilon_{0:k-1} \qquad (14) \\ &\propto w'_{k-1} \times \frac{p(\psi'_k \mid \psi'_{k-1})}{q(\psi'_k \mid \psi'_{k-1}, \mathbf{x}_k)} \int p(\mathbf{x}_k \mid \psi_k) p(\Upsilon_k) d\Upsilon_k \qquad (15) \\ &\propto w'_{k-1} \times \frac{p(\mathbf{x}_k \mid \psi'_k) p(\psi'_k \mid \psi'_{k-1})}{q(\psi'_k \mid \psi'_{k-1}, \mathbf{x}_k)} \qquad (16) \end{aligned}$$

The integral over $d\Upsilon_k$ in equation 15 is easily performed for the GLM when the likelihood function resembles equation 5 [7].

REFERENCES

[1] J. A. Simmons, M. B. Fenton, and M. J. O'Farrell, "Echolocation and pursuit of prey by bats," *Science*, vol. 203, Jan 1979.

[2] C. Dubois, M. Davy, and J. Idier, "Tracking of time-frequency components using particle filtering," *Proceedings (ICASSP '05) IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, Mar 2005.

[3] C. Andrieu, M. Davy, and A. Doucet, "Efficient particle filtering for jump Markov systems. Application to time-varying autoregressions," *IEEE Transactions on Signal Processing*, vol. 51, no. 7, Jul 2003.

[4] C. Dubois and M. Davy, "Joint detection and tracking of time-varying harmonic components: A flexible Bayesian approach," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 4, May 2007.

[5] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House Publishers, 2004.

[6] G. L. Bretthorst, *Bayesian Spectrum Analysis and Parameter Estimation*, Lecture Notes in Statistics. Springer-Verlag, 1988.

[7] J. J. K. ÓRuanaidh and W. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*, Springer, 1996.

[8] P. J. Green, "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, 1995.

[9] C. Andrieu and A. Doucet, "Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC," *IEEE Transactions on Signal Processing*, vol. 47, no. 10, Oct 1999.

[10] A. Jasra, A. Doucet, D. A. Stephens, and C. C. Holmes, "Interacting sequential Monte Carlo samplers for trans-dimensional simulation.," *Computational Statistics & Data Analysis*, vol. 52, no. 4, 2008.

[11] J.-R. Larocque, J. P. Reilly, and W. Ng, "Particle filters for tracking an unknown number of sources," *IEEE Transactions on Signal Processing*, vol. 50, no. 12, Dec 2002.

Proceedings of the Institute of Acoustics

A SEQUENTIAL MONTE CARLO ANALYSIS METHOD FOR BAT BIOSONAR SIGNALS

S Nagappa
J R Hopgood Institute for Digital Communications, University of Edinburgh, Edinburgh, UK

1 INTRODUCTION

Echolocation call structure is a key feature that provides a primary indication of bat species, specifically in cases where physical observation of the bat is not immediately possible or is difficult. Additionally, information on the structure of these calls can find application in other areas, for example, construction of biomedical or sonar signals.

In spite of several advances in frequency estimation techniques, the analysis of bat calls is still primarily performed using the short time Fourier transform (STFT) or wavelets [1]. The STFT estimates the frequency content of a signal by decomposing the signal onto a set of sinusoidal basis functions. The frequency content of the signal is evaluated within a sliding window, and as the window slides across the signal, a time-frequency representation of the signal is obtained.

Wavelet methods offer an alternative method for analysing a time-varying signal. A particular wavelet basis function is used to decompose the signal rather than the sinusoidal basis of the STFT. While the STFT uses a fixed window length, wavelets use a variable window length, with a smaller window used to detect features which occur on smaller scales. This use of a variable window size allows wavelet methods to improve on the limited time-frequency resolution of the STFT [1].

Both the STFT and wavelet methods are non-parametric methods and make no assumptions about the signal structure. Consequently, they rely on characterising the signal using a maximal number of basis functions. In the case of the STFT, for example, this corresponds to the large number of frequency bins in the Fourier transform. However, it is not necessary that any or all of these frequencies correspond to tones present in the signal under consideration.

In some instances, the STFT (or wavelets) is used as a preprocessing step for the bat calls. Using the spectrogram of the signal, other characteristics, for example, tail frequency or number of harmonics, can be determined. However, any further analysis is dependent on the results of the STFT operation. More complicated analysis of a call requires extracting the call from the time-frequency representation which may require other algorithms and assumptions.

It is possible to improve on the frequency estimates of the STFT by using a parametric approach which takes advantage of signal structure. For example, it is possible to model a bat call as an amplitude and frequency modulated tone containing harmonics. Such an approach is adopted here where a parametric frequency estimation algorithm [2] is applied to the analysis of bat echolocation calls.

The frequency estimation algorithm decomposes the signal into a small set of frequency modulated tones. In doing so, a ready estimate of the number of tones and their frequencies present in the signal is obtained, thus facilitating further signal analysis. The algorithm is implemented within a sequential framework, with the frequency content of the signal estimated within a sliding window. Sequential Monte Carlo (SMC) or particle filtering methods [3] are used to solve this problem within a Bayesian context.

The frequency estimation algorithm produces an estimate of the frequency content of the signal. These frequency estimates are then processed using a data association algorithm to form a coherent call which can be further analysed. Together, these two steps constitute the frequency tracker algorithm. The following sections will provide an overview of the frequency tracker algorithm as well as

results from its application to bat calls. A complete mathematical treatment of the algorithm as well as references to related methods is available in [2].

2 A PARTICLE FILTER FREQUENCY TRACKER

This section begins by examining the signal model used for bat calls. The signal model within the window is considered first. The frequency estimation algorithm estimates the frequencies within a sliding window of length L, L odd. Within this window, the signal is decomposed on to a variable number of chirp basis functions corresponding to the number of tones present in the signal. The use of a chirp basis as opposed to a sinusoidal basis is preferred when analysing frequency modulated signals [2]. Additionally, each chirp basis can have a number of harmonics as is used in the harmonic model described in [4].

Each instance of the evaluated window is termed a frame. The frame index is k and the frame is centred about the time $t_k \in \{0, \ldots, T-1\}$, where T is the length of the signal. The frequency content of the k^{th} frame is evaluated over the segment $l_k \in \{t_k - L', \ldots, t_k + L'\}$, where $L' = (L-1)/2$. For a signal containing P_k fundamental tones in the k^{th} frame, the observations can be written as:

$$x_{l_k} = \sum_{p=1}^{P_k} \sum_{h=1}^{H_k^{(p)}} \left(a_k^{(p,h)} \cos \phi_{k,l_k-t_k}^{(p,h)} + b_k^{(p,h)} \sin \phi_{k,l_k-t_k}^{(p,h)} \right) + n_{l_k} \tag{1}$$

where

$$\phi_{k,l}^{(p,h)} = 2\pi h \left(f_k^{(p)} + \frac{c_k^{(p)}}{2}(l-L) \right) l \tag{2}$$

In equations 1 and 2, x_{l_k} denotes the observed sample, $f_k^{(p)}$ is the frequency of the p^{th} tone and $H_k^{(p)}$ is the number of overtones associated with the p^{th} tone. The amplitudes of the cosine and sine contributions of the h^{th} harmonic of the p^{th} tone are respectively given by $a_k^{(p,h)}$ and $b_k^{(p,h)}$, and $n_{l_k} \sim \mathcal{N}(0, \sigma_{n,k}^2)$ is zero-mean white Gaussian observation noise with variance $\sigma_{n,k}^2$.

The algorithm is required to determine the parameters of the model described in equation 1. The estimation problem then reduces to estimating the smallest number of basis functions and their parameters that adequately represent the signal within the window. The estimated value of $f_k^{(p)}$ provides an estimate of the frequency of the p^{th} tone in the signal. The time-varying nature of these parameters indicates that the basis functions change over time, which allows the basis functions to deal with variations in the signal over time. The time-varying update is based on the assumption of a smoothly varying signal and enforces continuity between subsequent window frames.

Due to the nonlinear nature of the observation model, a particle filter [3] is adopted for sequential estimation of the parameters. Particle filters work by obtaining a sequential estimate of the posterior distribution on the parameters of interest. A number of samples, or particles, are used to form an adaptive non-uniformly sampled grid on which the posterior distribution is evaluated. This estimate of the posterior distribution can then be used to obtain estimates of the model parameters. The reader is once again referred to [2] for complete details of the algorithm.

The frequency estimation algorithm produces a discrete set of frequency estimates in the time-frequency domain. However, these discrete points, which may be likened to the spectrogram, do not provide an estimate of the echolocation call. It is necessary to link the points together to obtain an estimate of the call itself. This task is easily performed using data association algorithms and the algorithm chosen here is the nearest neighbour (NN) algorithm [5]. The next section will examine the application of the frequency tracker algorithm to bat echolocation calls.

3 ANALYSIS OF BAT CALLS

This section illustrates the application of the frequency tracker algorithm to field recordings of bat calls. In stark contrast with synthetic test signals, field data rarely satisfies all the assumptions made in the

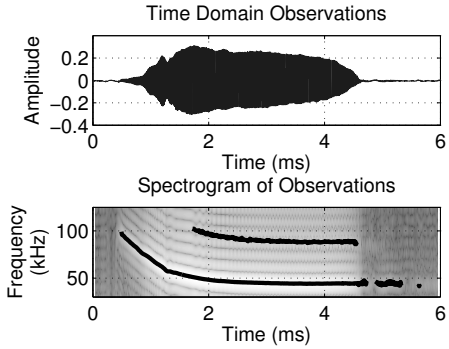


Figure 1: Analysis of a call from *P. pipistrellus*.

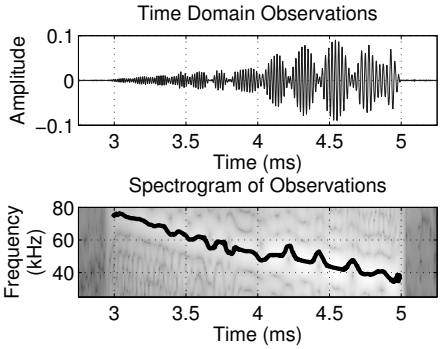


Figure 2: Analysis of a call from *M. daubentonii* shows the presence of notches in the call.

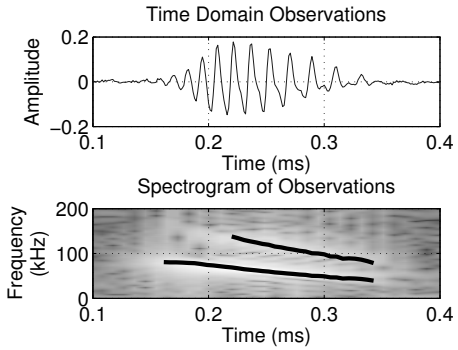


Figure 3: Analysis of a feeding buzz call from a pipistrelle.

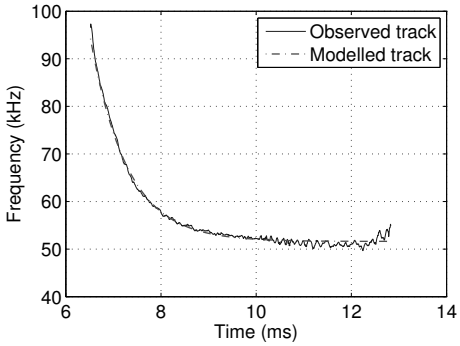


Figure 4: A call from *P. pygmaeus* is fit to an exponential frequency chirp.

model. In order to use the frequency tracker algorithm, it may be necessary to tune the algorithm as is briefly discussed in [2].

Results are presented from the analysis of calls from *Pipistrellus pipistrellus* (Schreber, 1774), *P. pygmaeus* (Leach, 1825) and *Myotis daubentonii* (Kuhl, 1817). The frequency tracker algorithm will be applied to calls under varying circumstances and shown to be both robust and flexible.

The frequency estimation results are presented as a set of extracted frequency calls overlaid on the STFT of the signal. Figure 1 shows an estimated call from *P. pipistrellus*. The harmonic model is able to capture the fundamental tone as well as the harmonic showing that the model provides a reasonable approximation to the real data. Figure 2 shows the analysis of a call from *M. daubentonii*. The frequency track shows the presence of upward and downward notches which are unclear in the STFT. Characterisation of such detail is one of the benefits of the flexible chirp model.

The advantages of the frequency tracker may also be observed when processing very short duration calls. Figure 3 shows a call from the feeding buzz of *P. pygmaeus*. At only 0.2 ms duration (150 samples when recorded at 750 kHz), the short duration of the call results in a large variance in the STFT time-frequency representation. In contrast, the more flexible frequency tracker method is able to easily extract the call as shown in the figure.

The extraction of calls simplifies the process of further analysis of the calls. The frequency tracker, for example, can be used to automate the call extraction process allowing the researcher to search for trends in calls.

The final section of these results shows how chirp model estimation can be applied to an extracted call. In the example here, a call from *P. pygmaeus* is extracted and tested against different chirp

models [6]. For each chirp model, the parameters are determined by obtaining the values which minimise the error between the estimated call and the chosen chirp model. The chirp model is then determined according to the goodness of fit between the estimated call and the different chirp models by selecting the model which provides the minimum error.

Figure 4 shows how an exponential chirp model provides a close approximation to the estimated call. This approach can be extended to the analysis of calls from several different species under varying circumstances to search for trends within an individual or species, or even across species. Further, by reducing a call to a model and a set of parameters, it may be possible to perform classification using a small set of parameters, rather than from examination of the spectrogram of a call.

4 CONCLUSIONS

The work presented here provides a method for improving the analysis of bat echolocation calls. The commonly used STFT has several limitations that can be overcome through the use of a parametric approach as used in the method described.

The application of the method shows that it provides a simple approach to the analysis of bat echolocation calls by providing a ready estimate of the frequency content of the call. The method is capable of estimating fine structure present in the signal and is shown to perform well with extremely short duration signals. The reduction of a call to a chirp model-type and a small set of parameters may permit a more quantitative analysis of signal structure as well as aid in classification tasks. The methods described here provide the scope for a broader analysis of bat echolocation signals and this is the subject of ongoing work.

ACKNOWLEDGEMENTS

IDCOM is in the Joint Research Institute for Signal and Image Processing, a member of the Edinburgh Research Partnership in Engineering and Mathematics. The work carried out here is sponsored by the EPSRC under the basic technology grant “Biologically Inspired Acoustic Systems” (EP/C523776/1). Bat recordings used in this work were obtained from Dr Dean Waters at the University of Leeds.

REFERENCES

1. P. Flandrin. *Time-Frequency/Time-scale Analysis (Wavelet Analysis and Its Applications)*. Academic Press, 1st edition, 1999.

2. S. Nagappa and J. R. Hopgood. Frequency tracking using monte carlo methods: Application to bat echolocation signals. In *Proceedings of the 16th European Signal Processing Conference (EUSIPCO 2008)*, 2008.

3. B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman Filter*. Artech House, 2004.

4. M. Davy, S. J. Godsill, and J. Idier. Bayesian analysis of polyphonic western tonal music. *Journal of the Acoustical Society of America (JASA)*, 119: pp. 2498–2517, 2006.

5. S. Blackman and R. Popoli. *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.

6. W. M. Masters, S. C. Jacobs, and J. A. Simmons. The structure of echolocation sounds used by the big brown bat *Eptesicus fuscus*: Some consequences for echo processing. *The Journal of the Acoustical Society of America*, 89(3): pp. 1402–1413, March 1991.

References

[1] J. A. Thomas, C. F. Moss, and M. Vater, Eds., *Echolocation in Bats and Dolphins*. The University of Chicago Press, 2004.

[2] A. M. Boonman, S. Parsons, and G. Jones, "The Influence of Flight Speed on the Ranging Performance of Bats using Frequency Modulated Echolocation Pulses," *Journal of the Acoustical Society of America (JASA)*, vol. 113, no. 1, pp. 617–628, January 2003.

[3] A. Boonman and H.-U. Schnitzler, "Frequency Modulation Patterns in the Echolocation Signals of Two Verspertilionid Bats," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 191, no. 1, pp. 13–21, January 2005.

[4] A. Boonman and J. Ostwald, "A Modeling Approach to Explain Pulse Design in Bats," *Biological Cybernetics*, vol. 97, no. 2, pp. 159–172, Aug 2007.

[5] J. A. Simmons, M. B. Fenton, and M. J. O. Farrell, "Echolocation and Pursuit of Prey by Bats," *Science*, vol. 203, no. 4375, pp. 16–21, 1979. [Online]. Available: http://www.jstor.org/stable/1747535

[6] Y. Kopsinis, E. Aboutanios, D. A. Waters, and S. McLaughlin, "Time-Frequency and Advanced Frequency Estimation Techniques for the Investigation of Bat Echolocation Calls," *Journal of the Acoustical Society of America (JASA)*, vol. 127, pp. 1124–1134, Feb 2010.

[7] L. Cohen, *Time-Frequency Analysis*. Prentice Hall, 1995.

[8] P. Flandrin, *Time-Frequency/Time-Scale Analysis (Wavelet Analysis and Its Applications)*, 1st ed. Academic Press, 1999.

[9] B. Boashash, "Estimating and Interpreting the Instantaneous Frequency of a Signal. I. Fundamentals," *Proc. IEEE*, vol. 80, no. 4, pp. 520–538, 1992.

[10] ——, "Estimating and Interpreting the Instantaneous Frequency of a Signal. II. Algorithms and applications," *Proc. IEEE*, vol. 80, no. 4, pp. 540–568, 1992.

[11] P. J. Loughlin and B. Tacer, "Comments on the Interpretation of Instantaneous Frequency," *IEEE Signal Process. Lett.*, vol. 4, no. 5, pp. 123–125, 1997.

[12] P. M. Oliveira and V. Barroso, "Instantaneous Frequency of Multicomponent Signals," *IEEE Signal Process. Lett.*, vol. 6, no. 4, pp. 81–83, 1999.

[13] P. J. Loughlin and B. Tacer, "Instantaneous Frequency and the Conditional Mean Frequency of a Signal," *Signal Processing*, vol. 60, no. 2, pp. 153–162, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/B6V18-46XGN35-2/2/285bf2a946118ff60c373c9adf9b487c

[14] P. J. Loughlin, "Spectrographic Measurement of Instantaneous Frequency and the Time-Dependent Weighted Average Instantaneous Frequency," *The Journal of the Acoustical Society of America*, vol. 105, no. 1, pp. 264–274, 1999. [Online]. Available: http://link.aip.org/link/?JAS/105/264/1

[15] M. I. Skolnik, *Introduction to Radar Systems*. McGraw-Hill Book Co., 1980.

[16] P. Flandrin, "Time Frequency and Chirps," H. H. Szu, D. L. Donoho, A. W. Lohmann, W. J. Campbell, and J. R. Buss, Eds., vol. 4391, no. 1. SPIE, 2001, pp. 161–175. [Online]. Available: http://link.aip.org/link/?PSI/4391/161/1

[17] G. L. Bretthorst, *Bayesian Spectrum Analysis and Parameter Estimation*, ser. Lecture Notes in Statistics. New York: Springer-Verlag, 1988.

[18] E. T. Jaynes, "Bayesian Spectrum and Chirp Analysis," in *Maximum-Entropy and Bayesian Spectral Analysis and Estimation Problems*, C. R. Smith and G. J. Erickson, Eds. D. Reidel Publishing Co., 1987.

[19] V. Katkovnik, "A New Form of the Fourier Transform for Time-Varying Frequency Estimation," in *Proc. URSI International Symposium on Signals, Systems, and Electronics ISSSE '95*, 1995, pp. 179–182.

[20] F. Auger and P. Flandrin, "Improving the Readability of Time-Frequency and Time-Scale Representations by the Reassignment Method," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1068–1089, 1995.

[21] X. Li and G. Bi, "A New Reassigned Time-Frequency Representation," in *Proceedings of the 16th European Signal Processing Conference (EUSIPCO 2008)*, 2008.

[22] S. W. Hainsworth, M. D. Macleod, and P. J. Wolfe, "Analysis of Reassigned Spectrograms for Musical Transcription," in *Proc. IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, 2001, pp. 23–26.

[23] F. Plante, G. Meyer, and W. A. Ainsworth, "Improvement of Speech Spectrogram Accuracy by the Method of Reassignment," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 3, pp. 282–287, 1998.

[24] P. Duvaut, A. Doucet, C. Veaux, and P. Flandrin, "Instantaneous Frequency Estimation: Bayesian Approaches Versus Reassignment-Application to Gravitational Waves," in *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on*, vol. 5, 7-10 May 1996, pp. 2968–2971.

[25] C. K. Chui, *An Introduction to Wavelets*. Academic Press, 1992.

[26] M. Holschneider, *Wavelets: An Analysis Tool*. Oxford University Press, 1995.

[27] M. A. Pinsky, *Introduction to Fourier Analysis and Wavelets*, J. Paul J. Sally, Ed. Brooks/Cole, 2002.

[28] C. Andrieu, A. Doucet, and P. Duvant, "Joint Bayesian Detection and Estimation of Sinusoids Embedded in Noise," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, A. Doucet, Ed., vol. 4, 1998, pp. 2245–2248.

[29] C. Andrieu and A. Doucet, "Joint Bayesian Model Selection and Estimation of Noisy Sinusoids via Reversible Jump MCMC," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2667–2676, Oct 1999.

[30] P. J. Green, "Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, Dec 1995.

[31] P. Djuric, S. Godsill, W. Fitzgerald, and P. Rayner, "Detection and Estimation of Signals by Reversible Jump Markov Chain Monte Carlo Computations," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, S. Godsill, Ed., vol. 4, 1998, pp. 2269–2272 vol.4.

[32] C. Andrieu, N. De Freitas, and A. Doucet, "Sequential MCMC for Bayesian Model Selection," in *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics*, N. De Freitas, Ed., 1999, pp. 130–134.

[33] K. Copsey, N. Gordon, and A. Marrs, "Bayesian Analysis of Generalized Frequency-Modulated Signals," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 725–735, 2002.

[34] J.-R. Larocque, J. Reilly, and W. Ng, "Particle Filters for Tracking an Unknown Number of Sources," *IEEE Trans. Signal Process.*, vol. 50, no. 12, pp. 2926–2937, 2002.

[35] J.-R. Larocque and J. Reilly, "Reversible Jump MCMC for Joint Detection and Estimation of Sources in Colored Noise," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 231–240, 2002.

[36] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed., ser. Springer Texts in Statistics. Springer, 2004.

[37] J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*. Wiley, 1994.

[38] M. Davy, S. J. Godsill, and J. Idier, "Bayesian Analysis of Polyphonic Western Tonal Music," *Journal of the Acoustical Society of America (JASA)*, vol. 119, pp. 2498–2517, 2006.

[39] J. Makhoul, "Linear Prediction: A Tutorial Review," *Proc. IEEE*, vol. 63, no. 4, pp. 561–580, 1975.

[40] ——, "Spectral Analysis of Speech by Linear Prediction," *IEEE Trans. Audio Electroacoust.*, vol. 21, no. 3, pp. 140–148, 1973.

[41] ——, "Spectral Linear Prediction: Properties and Applications," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 23, no. 3, pp. 283–296, 1975.

[42] K. Schnell and A. Lacroix, "Time-Varying Linear Prediction for Speech Analysis and Aynthesis," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP 2008*, 2008, pp. 3941–3944.

[43] B. F. La Scala and R. R. Bitmead, "Design of an Extended Kalman Filter Frequency Tracker," *IEEE Trans. Signal Process.*, vol. 44, no. 3, pp. 739–742, 1996.

[44] B. F. La Scala, R. R. Bitmead, and B. G. Quinn, "An Extended Kalman Filter Frequency Tracker for High-Noise Environments," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 431–434, 1996.

[45] H. Hajimolahoseini, M. R. Taban, and H. R. Abutalebi, "Improvement of Extended Kalman Filter Frequency Tracker for Nonstationary Harmonic Signals," in *Proc. International Symposium on Telecommunications IST 2008*, 2008, pp. 592–597.

[46] R. L. Streit and R. F. Barrett, "Frequency Line Tracking using Hidden Markov Models," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 4, pp. 586–598, 1990.

[47] R. A. Carmona, W. L. Hwang, and B. Torresani, "Multiridge Detection and Time-Frequency Reconstruction," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 480–492, 1999.

[48] C. Dubois, M. Davy, and J. Idier, "Tracking of Time-Frequency Components using Particle Filtering," in *Proceedings (ICASSP '05) IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, March 2005.

[49] C. Dubois and M. Davy, "Harmonic Tracking using Sequential Monte Carlo," in *Proc. IEEE/SP 13th Workshop on Statistical Signal Processing*, 2005, pp. 1292–1297.

[50] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.

[51] D. Clark, A.-T. Cemgil, P. Peeling, and S. Godsill, "Multi-Object Tracking of Sinusoidal Components in Audio with the Gaussian Mixture Probability Hypothesis Density Filter," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, A.-T. Cemgil, Ed., 2007, pp. 339–342.

[52] D. Clark, A. T. Cemgil, P. H. Peeling, and S. J. Godsill, "Multi-Object Tracking of Sinusoidal Components in Audio with the Gaussian Mixture Probability Hypothesis Density Filter," Department of Engineering, University of Cambridge, Tech. Rep., 2007.

[53] K. Panta, B. Vo, and S. Singh, "Improved Probability Hypothesis Density (PHD) Filter for Multitarget Tracking," in *Proc. Third International Conference on Intelligent Sensing and Information Processing ICISIP 2005*, 2005, pp. 213–218.

[54] B.-N. Vo and W.-K. Ma, "The Gaussian Mixture Probability Hypothesis Density Filter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4091–4104, 2006.

[55] D. Clark, K. Panta, and B.-N. Vo, "The GM-PHD Filter Multiple Target Tracker," in *Proc. 9th International Conference on Information Fusion*, K. Panta, Ed., 2006, pp. 1–8.

[56] P. Djuric and S. Kay, "Parameter Estimation of Chirp Signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 12, pp. 2118 – 2126, Dec 1990.

[57] R. M. Liang and K. S. Arun, "Parameter Estimation for Superimposed Chirp Signals," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP-92*, vol. 5, 1992, pp. 273–276.

[58] S. Saha and S. Kay, "Maximum Likelihood Parameter Estimation of Superimposed Chirps using Monte Carlo Importance Sampling," *IEEE Transactions of Signal Processing*, vol. 50, no. 2, pp. 224–230, Feb 2002.

[59] O. Besson, N. Ghogho, and A. Swami, "Parameter Estimation for Random Amplitude Chirp Signals," *IEEE Trans. Signal Process.*, vol. 47, no. 12, pp. 3208–3219, 1999.

[60] S. Tretter, "Estimating the Frequency of a Noisy Sinusoid by Linear Regression (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 31, no. 6, pp. 832–835, 1985.

[61] S. Kay and S. Saha, "Mean Likelihood Frequency Estimation," *IEEE Trans. Signal Process.*, vol. 48, no. 7, pp. 1937–1946, 2000.

[62] J. J. K. ÓRuanaidh and W. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*. Springer, 1996.

[63] W. W. L. Au, "Comparison of Bat and Dolphin Sonar Capabilities," in *Echolocation in Bats and Dolphins*, J. A. Thomas, C. F. Moss, and M. Vater, Eds. The University of Chicago Press, 2004, pp. xiii–xxvii.

[64] A. Denzinger, E. K. V. Kalko, and G. Jones, "Ecological and Evolutionary Aspects of Echolocation in Bats," in *Echolocation in Bats and Dolphins*, J. A. Thomas, C. F. Moss, and M. Vater, Eds. The University of Chicago Press, 2004, pp. 311–326.

[65] C. Karine and E. Kalko, "Toward a Global Bat-Signal Database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 81–85, 2001.

[66] R. A. Altes and E. L. Titlebaum, "Bat Signals as Optimally Doppler Tolerant Waveforms," *The Journal of the Acoustical Society of America*, vol. 48, no. 4, pp. 1014–1020, 1970.

[67] W. M. Masters, S. C. Jacobs, and J. A. Simmons, "The Structure of Echolocation Sounds used by the Big Brown Bat *Eptesicus fuscus*: Some consequences for Echo Processing," *The Journal of the Acoustical Society of America*, vol. 89, no. 3, pp. 1402–1413, March 1991.

[68] W. M. Masters and K. A. S. Raver, "Bats Learning to Use Echoes with Unfamiliar Time-Frequency Structures," in *Echolocation in Bats and Dolphins*, J. A. Thomas, C. F. Moss, and M. Vater, Eds. The University of Chicago Press, 2004, pp. 273–277.

[69] G. Neuweiler and E. Covey, *The Biology of Bats*. Oxford University Press USA, 2000.

[70] C. Dubois and M. Davy, "Joint Detection and Tracking of Time-Varying Harmonic Components: A Flexible Bayesian Approach," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 4, pp. 1283–1295, May 2007.

[71] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and Adaptive Signal Processing*. McGraw-Hill Book Co., 2000.

[72] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter*. Artech House, 2004.

[73] S. J. Julier and J. K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems," in *Proc. Areosense: The 11th International Symposium on Aerospace/Defence Sensing, Simulation, and Controls*, 1997.

[74] E. A. Wan and R. V. D. Merwe, "The Unscented Kalman Filter," in *Kalman Filtering and Neural Networks*, S. Haykin, Ed. John Wiley & Sons Inc, 2001, pp. 221–280.

[75] M. K. Pitt and N. Shephard, "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, vol. 94, no. 446, pp. 590–599, Jun 1999.

[76] A. Doucet, S. Godsill, and C. Andrieu, "On Sequential Monte Carlo Sampling Methods for Byesian Fltering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[77] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, 2001.

[78] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb 2002.

[79] A. Doucet, N. Gordon, and V. Krishnamurthy, "Particle Filters for State Estimation of Jump Markov Linear Systems," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 613–624, 2001.

[80] R. Douc, O. Cappé, and E. Moulines, "Comparison of Resampling Schemes for Particle Filtering," in *4th International Symposium on Image and Signal Processing and Analysis (ISPA)*, 2005, pp. 64–69.

[81] J. D. Hol, T. B. Schon, and F. Gustafsson, "On Resampling Algorithms for Particle Filters," in *Proc. IEEE Nonlinear Statistical Signal Processing Workshop*, 2006, pp. 79–82.

[82] A. Doucet, N. de Freitas, K. Murphy, and S. Russell, "Rao-Blackwellised Particle Filtering for Dynamic Bayesian Networks," in *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence*, 2000, pp. 176–183.

[83] F. Mustiere, M. Bolic, and M. Bouchard, "Rao-Blackwellised Particle Filters: Examples of Applications," in *Proc. Canadian Conference on Electrical and Computer Engineering CCECE '06*, 2006, pp. 1196–1200.

[84] K. Murphy and S. Russell, "Rao-Blackwellised Particle Filtering for Dynamic Bayesian Networks," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. Gordon, Eds. Springer, 2001, pp. 499–515.

[85] N. de Freitas, "Rao-Blackwellised Particle Filtering for Fault Diagnosis," in *Proc. IEEE Aerospace*, vol. 4, 2002, pp. 1767–1772.

[86] S. Godsill and J. Vermaak, "Models and Algorithms for Tracking using Trans-Dimensional Sequential Monte Carlo," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, J. Vermaak, Ed., vol. 3, 2004, pp. 976–979.

[87] A. Ooi, B.-N. Vo, and A. Doucet, "Particle Filtering for Multi-Target Tracking using Jump Markov Systems," in *Proc. Intelligent Sensors, Sensor Networks and Information Processing Conference the 2004*, B.-N. Vo, Ed., 2004, pp. 131–136.

[88] C. Andrieu, M. Davy, and A. Doucet, "Efficient Particle Filtering for Jump Markov Systems," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '02)*, M. Davy, Ed., vol. 2, 2002, pp. 1625–1628.

[89] H. Driessen and Y. Boers, "Efficient Particle Filter for Jump Markov Nonlinear Systems," *IEE Proceedings -Radar, Sonar and Navigation*, vol. 152, no. 5, pp. 323–326, 2005.

[90] A. Jasra, A. Doucet, D. A. Stephens, and C. C. Holmes, "Interacting Sequential Monte Carlo Samplers for Trans-Dimensional Simulation," *Computational Statistics & Data Analysis*, vol. 52, no. 4, pp. 1765–1791, 2008.

[91] C. Andrieu, M. Davy, and A. Doucet, "Improved Auxiliary Particle Filtering: Applications to Time-Varying Spectral Analysis," in *Proc. 11th IEEE Signal Processing Workshop on Statistical Signal Processing*, M. Davy, Ed., 2001, pp. 309–312.

[92] ——, "Efficient Particle Filtering for Jump Markov Systems. Application to Time-Varying Autoregressions," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1762–1770, 2003.

[93] R. van der Merwe, N. de Freitas, A. Doucet, and E. Wan, "The Unscented Particle Filter," in *Advances in Neural Information Processing Systems 13*, Nov 2001. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.32.9011

[94] Y. Li, A. Papandreou-Suppappola, and D. Morrell, "Instantaneous Frequency Estimation using Sequential Bayesian Techniques," in *Proc. Fortieth Asilomar Conference on Signals, Systems and Computers ACSSC '06*, 2006, pp. 569–573.

[95] D. Lee and N. Chia, "A Particle Algorithm for Sequential Bayesian Parameter Estimation and Model Selection," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 326–336, 2002.

[96] N. Chopin, "A Sequential Particle Filter Method for Static Models," *Biometrika*, vol. 89, no. 3, pp. 539–552, 2002. [Online]. Available: http://biomet.oxfordjournals.org/cgi/content/abstract/89/3/539

[97] S. Kullback and R. A. Leibler, "On Information and Sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[98] M. K. Obrist, R. Boesch, P. F. Flückiger, and U. Dieckmann, "Who's Calling? Acoustic Bat Species Identification Revised with Synergetics," in *Echolocation in Bats and Dolphins*, J. A. Thomas, C. F. Moss, and M. Vater, Eds. The University of Chicago Press, 2004, pp. 484–492.

[99] S. Peleg and B. Porat, "The Cramer-Rao Lower Bound for Signals with Constant Amplitude and Polynomial Phase," *Signal Processing, IEEE Transactions on*, vol. 39, no. 3, pp. 749–752, Mar 1991.

[100] B. Ristic and B. Boashash, "Comments on The Cramer-Rao Lower Bounds for Signals with Constant Amplitude and Polynomial Phase," *Signal Processing, IEEE Transactions on*, vol. 46, no. 6, pp. 1708–1709, Jun 1998.

[101] M. Davy and S. Godsill, "Bayesian Harmonic Models for Musical Signal Analysis," in *Bayesian Statistics VII.*, J. Bernardo, Ed. Oxford University Press, 2002. [Online]. Available: citeseer.ist.psu.edu/davy02bayesian.html

[102] G. Storvik, "Particle Filters for State-Space Models with the Presence of Unknown Static Parameters," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 281–289, Feb 2002.

[103] J. R. Hopgood, "Statistical Signal Processing," Lecture Notes, University of Edinburgh, 2007.

[104] H. Driessen and Y. Boers, "MAP Estimation in Particle Filter Tracking," in *Proc. IET Seminar on Target Tracking and Data Fusion: Algorithms and Applications*, 2008, pp. 41–45.

[105] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. The MIT Press, 2001.

[106] C. K. Yap, *Fundamental Problems of Algorithmic Algebra*. Oxford University Press, 1999.

[107] M. B. Fenton, "Aerial-Feeding Bats: Getting the Most Out of Echolocation," in *Echolocation in Bats and Dolphins*, J. A. Thomas, C. F. Moss, and M. Vater, Eds. The University of Chicago Press, 2004, pp. 350–355.

[108] D. L. Hall and J. Llinas, Eds., *Handbook of Multisensor Data Fusion*. CRC Press, 2001.